

# Degraded Broadcast Channel with Side Information, Confidential Messages and with or without Noiseless Feedback

Bin Dai, A. J. Han Vinck, *Fellow, IEEE*, Zhuojun Zhuang, and Yuan Luo

## Abstract

In this paper, first, we investigate the model of degraded broadcast channel with side information and confidential messages. This work is from Steinberg's work on the degraded broadcast channel with causal and noncausal side information, and Csiszár-Körner's work on broadcast channel with confidential messages. Inner and outer bounds on the capacity-equivocation regions are provided for the noncausal and causal cases. Superposition coding and double-binning technique are used in the corresponding achievability proofs.

Then, we investigate the degraded broadcast channel with side information, confidential messages and noiseless feedback. The noiseless feedback is from the non-degraded receiver to the channel encoder. Inner and outer bounds on the capacity-equivocation region are provided for the noncausal case, and the capacity-equivocation region is determined for the causal case. Compared with the model without feedback, we find that the noiseless feedback helps to enlarge the inner bounds for both causal and noncausal cases. In the achievability proof of the feedback model, the noiseless feedback is used as a secret key shared by the non-degraded receiver and the transmitter, and therefore, the code construction for the feedback model is a combination of superposition coding, Gel'fand-Pinsker's binning, block Markov coding and Ahlswede-Cai's secret key on the feedback system.

## Index Terms

Confidential message, capacity-equivocation region, degraded broadcast channel, noiseless feedback, secrecy capacity, side information.

## I. INTRODUCTION

Equivocation was first introduced into channel coding by Wyner in his study of wiretap channel [15]. It is a kind of degraded broadcast channels. The object is to transmit messages to the legitimate receiver, while keeping the

B. Dai is with the Computer Science and Engineering Department, Shanghai Jiao Tong University, and the Institute for Experimental Mathematics, Duisburg-Essen University, Ellernstr.29, 45326 Essen, Germany e-mail: daibin007@sjtu.edu.cn.

A. J. Han Vinck is with the Institute for Experimental Mathematics, Duisburg-Essen University, Ellernstr.29, 45326 Essen, Germany e-mail: vinck@iem.uni-due.de.

Z. Zhuang and Y. Luo are with the Computer Science and Engineering Department, Shanghai Jiao Tong University, Shanghai 200240, China e-mail: zhuojunzzj@sjtu.edu.cn, yuanluo@sjtu.edu.cn.

wiretapper as ignorant of the messages as possible. After the publication of Wyner's work, Csiszár and Körner [4] investigated a more general situation: the broadcast channels with confidential messages, see Figure 1. The model of [4] is to transmit confidential messages to receiver 1 at rate  $R_1$  and common messages to both receivers at rate  $R_0$ , while keeping receiver 2 as ignorant of the confidential messages as possible. Measuring ignorance by equivocation, a single-letter characterization of all the achievable triples  $(R_1, R_e, R_0)$  was provided in [4], where  $R_e$  is the second receiver's equivocation to the confidential messages. Note that the model of [4] is also a generalization of [8], where no confidentiality condition is imposed. In addition, Merhav [11] studied a specified wiretap channel, and obtained the capacity region, where both the legitimate receiver and the wiretapper have access to some leaked symbols from the source, but the channels for the wiretapper are more noisy than the legitimate receiver, which shares a secret key with the encoder.

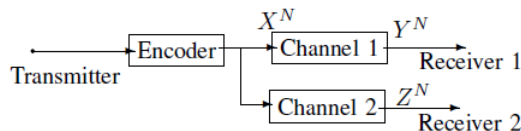


Fig. 1: Broadcast channels with confidential messages

In communication systems there is often a feedback link from the receiver to the transmitter, e.g. the two-way channels for telephone connections. It is well known that feedback does not increase the capacity of discrete memoryless channel (DMC). However, does the feedback increase the capacity region of the wiretap channel? In order to solve this problem, Ahlswede and Cai studied the general wiretap channel (the wiretap channel does not need to be degraded) with noiseless feedback from the legitimate receiver [1], and both upper and lower bounds of the secrecy capacity were provided. Specifically, for the degraded case, they showed that the secrecy capacity is larger than that of Wyner's wiretap channel (without feedback). In the achievability proof, Ahlswede and Cai [1] used the noiseless feedback as a secret key shared by the transmitter and the legitimate receiver, while the wiretapper had no additional knowledge about the key except his own received symbols. Besides Ahlswede and Cai's work, the wiretap channel with noisy feedback was studied in [10], and the wiretap channel with secure rate-limited feedback was studied in [2], and both of them focused on bounds of the secrecy capacity.

The coding for channels with causal side information at the encoder was first investigated by Shannon [13] in 1958. After that, in order to solve the problem of coding for a computer memory with defective cells, Kuznetsov and Tsybakov [9] considered a channel in the presence of noncausal side information at the transmitter. They provided some coding techniques without determination of the capacity. The capacity was found in 1980 by Gel'fand and Pinsker [7]. Furthermore, Costa [6] investigated a power-constrained additive noise channel, where part of the noise is known at the transmitter as side information. This channel is also called dirty paper channel. Based on the dirty paper channel, C. Mitropant et al. [12] studied the Gaussian wiretap channel with side information, and provided an inner bound on the capacity-equivocation region. Furthermore, Y. Chen et al. [3] investigated the

discrete memoryless wiretap channel with noncausal side information, and also provided an inner bound on the capacity-equivocation region. Note that the coding scheme of [3] is a combination of those in [7], [15]. In order to introduce side information to the broadcast channel, Steinberg investigated the degraded broadcast channel with side information [14], where both causal and noncausal side information were considered in his paper. Specifically, inner and outer bounds on capacity region were provided for the degraded broadcast channel with noncausal side information [14], and meanwhile, the capacity region of the degraded broadcast channel with causal side information was totally determined [14].

In this paper, we study the model of degraded broadcast channel with side information, confidential messages and with or without noiseless feedback (see Figure 2). The model of this paper is from Steinberg's work on degraded broadcast channel with side information at the encoder, and the model of broadcast channel with confidential messages provided by Csiszár and Körner. In Figure 2,  $S$  is the confidential message sent to receiver 1, and  $T$  is the common message sent to both receiver 1 and receiver 2. The transition probability of channel 1 depends on the side information  $V^N$ , and  $V^N$  is available to the channel encoder in a causal or noncausal manner. Receiver 2 can get a degraded version of  $Y^N$  via channel 2. In addition, there may exist a noiseless feedback from the output of channel 1 to the channel encoder.

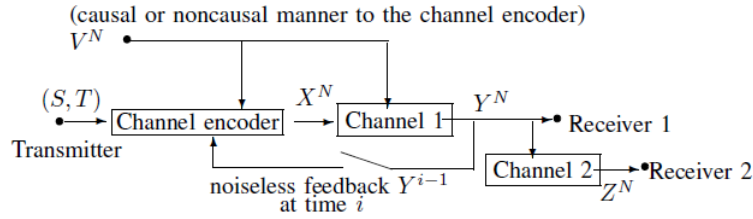


Fig. 2: Degraded broadcast channel with side information, confidential messages and with or without noiseless feedback

In this paper, random variables, sample values and alphabets are denoted by capital letters, lower case letters and calligraphic letters, respectively. A similar convention is applied to the random vectors and their sample values. **For example,  $U^N$  denotes a random  $N$ -vector  $(U_1, \dots, U_N)$ , and  $u^N = (u_1, \dots, u_N)$  is a specific vector value in  $\mathcal{U}^N$  that is the  $N$ th Cartesian power of  $\mathcal{U}$ .  $U_i^N$  denotes a random  $N - i + 1$ -vector  $(U_i, \dots, U_N)$ , and  $u_i^N = (u_i, \dots, u_N)$  is a specific vector value in  $\mathcal{U}_i^N$ .** Let  $p_V(v)$  denote the probability mass function  $Pr\{V = v\}$ . Throughout the paper, the logarithmic function is to the base 2.

The organization of this paper is as follows. In Section II, inner and outer bounds on the capacity-equivocation regions of the model of Figure 2 without feedback are provided by Theorem 1 to Theorem 4. Specifically, the inner and outer bounds for the model of Figure 2 with noncausal side information and without feedback are provided in Theorem 1 and Theorem 2, respectively. The inner and outer bounds for the model with causal side information and without feedback are provided in Theorem 3 and Theorem 4, respectively. In Section III, inner and outer bounds on

the capacity-equivocation region of the model of Figure 2 with noncausal side information and noiseless feedback are provided in Theorem 5 and Theorem 6, respectively. The capacity-equivocation region of the model with causal side information and noiseless feedback is given in Theorem 7. Section IV is for a binary example about the model of Figure 2 with causal side information and noiseless feedback. Final conclusions are provided in Section V.

## II. THE MODEL OF FIGURE 2 WITHOUT FEEDBACK

In this section, the model of Figure 2 without feedback is considered into two parts. The model of Figure 2 with noncausal side information and without feedback is described in Subsection II-A, and the model of Figure 2 with causal side information and without feedback is described in Subsection II-B, see the following.

### A. The Model of Figure 2 with Noncausal Side Information and without Feedback

In this subsection, a description of the model of Figure 2 with noncausal side information and without feedback is given by Definition 1 to Definition 5. The inner and outer bounds on the capacity-equivocation region  $\mathcal{R}^{(n)}$  composed of all achievable  $(R_0, R_1, R_e)$  triples are in Theorem 1 and Theorem 2, where the achievable  $(R_0, R_1, R_e)$  triple is defined in Definition 6.

**Definition 1: (Channel encoder)** The confidential message  $S$  takes values in  $\mathcal{S}$ , and the common message  $T$  takes values in  $\mathcal{T}$ .  $S$  and  $T$  are independent and uniformly distributed over their ranges.  $V^N$  is the side information of channel 1, and it is the output of a discrete memoryless source  $p_V(v)$ . In addition,  $V^N$  is available to the channel encoder in a noncausal manner. Note that  $V^N$  is independent of  $S$  and  $T$ . At the  $i$ -th time, the inputs of the channel encoder are  $S$ ,  $T$  and  $V^N$ , while the output is  $X_i$ , i.e., the channel encoder is a mapping

$$f_i : \mathcal{S} \times \mathcal{T} \times \mathcal{V}^N \rightarrow \mathcal{X}_i, \quad (2.1)$$

where  $f_i(s, t, v^N) = x_i \in \mathcal{X}$ ,  $s \in \mathcal{S}$ ,  $t \in \mathcal{T}$  and  $v^N \in \mathcal{V}^N$ . The transmission rates of the confidential message and the common message are  $\frac{\log \|\mathcal{S}\|}{N}$  and  $\frac{\log \|\mathcal{T}\|}{N}$ , respectively.

**Definition 2: (Channel 1)** Channel 1 is a DMC with finite input alphabet  $\mathcal{X} \times \mathcal{V}$ , finite output alphabet  $\mathcal{Y}$ , and transition probability  $Q_1(y|x, v)$ , where  $x \in \mathcal{X}$ ,  $v \in \mathcal{V}$ ,  $y \in \mathcal{Y}$ .  $Q_1(y^N|x^N, v^N) = \prod_{n=1}^N Q_1(y_n|x_n, v_n)$ . The inputs of Channel 1 are  $X^N$  and  $V^N$ , while the output is  $Y^N$ .

**Definition 3: (Channel 2)** Channel 2 is a DMC with finite input alphabet  $\mathcal{Y}$ , finite output alphabet  $\mathcal{Z}$ , and transition probability  $Q_2(z|y)$ , where  $y \in \mathcal{Y}$ ,  $z \in \mathcal{Z}$ .  $Q_2(z^N|y^N) = \prod_{n=1}^N Q_2(z_n|y_n)$ . The inputs of Channel 2 is  $Y^N$ , while the output is  $Z^N$ . Receiver 2's equivocation to the confidential message is defined as

$$\Delta = \frac{1}{N} H(S|Z^N). \quad (2.2)$$

The cascade of Channel 1 and Channel 2 is another DMC with transition probability

$$Q_3(z|x, v) = \sum_{y \in \mathcal{Y}} Q_W(z|y) Q_M(y|x, v). \quad (2.3)$$

**Definition 4: (Decoder 1)** Decoder 1 is a mapping  $f_{D1} : \mathcal{Y}^N \rightarrow \mathcal{S} \times \mathcal{T}$ , with input  $Y^N$  and outputs  $\hat{S}, \hat{T}$ . Let  $P_{e1}$  be the error probability of receiver 1, and it is defined as  $Pr\{(S, T) \neq (\hat{S}, \hat{T})\}$ .

**Definition 5: (Decoder 2)** Decoder 2 is a mapping  $f_{D2} : \mathcal{Z}^N \rightarrow \mathcal{T}$ , with input  $Z^N$  and output  $\tilde{T}$ . Let  $P_{e2}$  be the error probability of receiver 2, and it is defined as  $Pr\{T \neq \tilde{T}\}$ .

**Definition 6: (Achievable  $(R_0, R_1, R_e)$  triple in the model of Figure 2 with noncausal side information and without feedback)** A triple  $(R_0, R_1, R_e)$  (where  $R_0, R_1, R_e > 0$ ) is called achievable if, for any  $\epsilon > 0$  (where  $\epsilon$  is an arbitrary small positive real number and  $\epsilon \rightarrow 0$ ), there exists a channel encoder-decoder  $(N, \Delta, P_{e1}, P_{e2})$  such that

$$\lim_{N \rightarrow \infty} \frac{\log \|\mathcal{T}\|}{N} = R_0, \lim_{N \rightarrow \infty} \frac{\log \|\mathcal{S}\|}{N} = R_1, \lim_{N \rightarrow \infty} \Delta \geq R_e, P_{e1} \leq \epsilon, P_{e2} \leq \epsilon. \quad (2.4)$$

The following Theorem 1 and Theorem 2 provide inner and outer bounds on the capacity-equivocation region  $\mathcal{R}^{(n)}$  for the model of Figure 2 with noncausal side information and without feedback, and they are proved in Appendix A and Appendix B.

**Theorem 1:** A single-letter characterization of the region  $\mathcal{R}^{(ni)}$  is as follows,

$$\begin{aligned} \mathcal{R}^{(ni)} = \{ & (R_0, R_1, R_e) : 0 \leq R_e \leq R_1, \\ & R_0 \leq I(U; Z) - I(U; V), \\ & R_1 \leq I(K; Y|U) - I(K; V|U), \\ & R_e \leq I(K; Y|U) - I(K; Z|U) \}, \end{aligned}$$

where  $p_{UKVXYZ}(u, k, v, x, y, z) = p_{Z|Y}(z|y)p_{Y|X,V}(y|x, v)p_{UKXV}(u, k, x, v)$ , which implies that  $(U, K) \rightarrow (X, V) \rightarrow Y \rightarrow Z$ .

The region  $\mathcal{R}^{(ni)}$  satisfies  $\mathcal{R}^{(ni)} \subseteq \mathcal{R}^{(n)}$ .

**Remark 1:** There are some notes on Theorem 1, see the following.

- The region  $\mathcal{R}^{(ni)}$  is convex, and the proof is directly obtained by introducing a time sharing random variable into Theorem 1, and therefore, we omit the proof here.
- The ranges of the random variables  $U$  and  $K$  satisfy

$$\begin{aligned} \|\mathcal{U}\| & \leq \|\mathcal{X}\| \|\mathcal{V}\| + 2, \\ \|\mathcal{K}\| & \leq (\|\mathcal{X}\| \|\mathcal{V}\| + 2)^2. \end{aligned}$$

The proof is in Appendix C.

- Without the secrecy parameter  $R_e$ , the region  $\mathcal{R}^{(ni)}$  is exactly the same as the achievable region for the degraded broadcast channel with noncausal side information [14].
- The points in  $\mathcal{R}^{(n)}$  for which  $R_e = R_1$  are of considerable interest, which imply  $H(S) = H(S|Z^N)$ .

**Definition 7: (The secrecy capacity  $C_s^{(n)}$ )** The secrecy capacity  $C_s^{(n)}$  of the model of Figure 2 with noncausal side information and without feedback, is denoted by

$$C_s^{(n)} = \max_{(R_0=0, R_1, R_e=R_1) \in \mathcal{R}^{(n)}} R_1. \quad (2.5)$$

Furthermore, the secrecy capacity  $C_s^{(n)}$  satisfies

$$C_s^{(n)} \geq \max \min \{I(K; Y|U) - I(K; Z|U), I(K; Y|U) - I(K; V|U)\}. \quad (2.6)$$

*Proof of (2.6):* Substituting  $R_e = R_1$  and  $R_0 = 0$  into the region  $\mathcal{R}^{(ni)}$  in Theorem 1, we have

$$R_1 \leq I(K; Y|U) - I(K; V|U), \quad (2.7)$$

$$R_1 \leq I(K; Y|U) - I(K; Z|U). \quad (2.8)$$

$$(2.9)$$

Note that the triple  $(R_0 = 0, R_1 = \max \min \{I(K; Y|U) - I(K; Z|U), I(K; Y|U) - I(K; V|U)\}, R_e = R_1)$  is achievable, and therefore, the secrecy capacity  $C_s^{(n)} \geq \max \min \{I(K; Y|U) - I(K; Z|U), I(K; Y|U) - I(K; V|U)\}$ . Thus the proof is completed. ■

*Theorem 2:* A single-letter characterization of the region  $\mathcal{R}^{(no)}$  is as follows,

$$\mathcal{R}^{(no)} = \{(R_0, R_1, R_e) : 0 \leq R_e \leq R_1,$$

$$R_0 \leq I(U; Z) - I(U; V),$$

$$R_1 \leq I(K; Y|U, A) - I(K; V|U, A),$$

$$R_0 + R_1 \leq I(U, A, K; Y) - I(U, A, K; V),$$

$$R_e \leq I(K; Y|U, A) - I(K; V|U, A) - I(K; Z|U) + I(K; V|U)\},$$

where  $p_{UKAVXYZ}(u, k, a, v, x, y, z) = p_{Z|Y}(z|y)p_{Y|X,V}(y|x, v)p_{UKAXV}(u, k, a, x, v)$ , which implies that  $(U, K, A) \rightarrow (X, V) \rightarrow Y \rightarrow Z$ .

The region  $\mathcal{R}^{(no)}$  satisfies  $\mathcal{R}^{(n)} \subseteq \mathcal{R}^{(no)}$ .

*Remark 2:* There are some notes on Theorem 2, see the following.

- The region  $\mathcal{R}^{(no)}$  is convex, and the proof is directly obtained by introducing a time sharing random variable into Theorem 2, and therefore, we omit the proof here.
- The ranges of the random variables  $U$  and  $K$  satisfy

$$\|\mathcal{U}\| \leq \|\mathcal{X}\| \|\mathcal{V}\| + 2,$$

$$\|\mathcal{A}\| \leq (\|\mathcal{X}\| \|\mathcal{V}\| + 2)(\|\mathcal{X}\| \|\mathcal{V}\| + 1),$$

$$\|\mathcal{K}\| \leq (\|\mathcal{X}\| \|\mathcal{V}\| + 2)(\|\mathcal{X}\| \|\mathcal{V}\| + 1)(\|\mathcal{X}\| \|\mathcal{V}\| + 2)^2.$$

The proof is similar to the proof of the second part of Remark 1, and it is omitted here.

- Without the secrecy parameter  $R_e$ , the region  $\mathcal{R}^{(no)}$  is exactly the same as the outer bound for the degraded broadcast channel with noncausal side information [14].

### B. The Model of Figure 2 with Causal Side Information and without Feedback

The model of Figure 2 with causal side information and without feedback is similar to the model in Subsection II-A, except that the side information  $V^N$  in Definition 1 is known to the channel encoder in a causal manner, i.e., at the  $i$ -th time ( $1 \leq i \leq N$ ), the output of the channel encoder  $x_i = f_i(s, t, v^i)$ , where  $v^i = (v_1, v_2, \dots, v_i)$  and  $f_i$  is the mapping at time  $i$ . Note that  $V_i$  is independent of  $(Y^{i-1}, T, V_{i+1}^N, Z^{i-1})$ , where  $V_{i+1}^N = (V_{i+1}, V_{i+2}, \dots, V_N)$ ,  $Y^{i-1} = (Y_1, Y_2, \dots, Y_{i-1})$  and  $Z^{i-1} = (Z_1, Z_2, \dots, Z_{i-1})$ .

The following Theorem 3 and Theorem 4 provide inner and outer bounds on the capacity-equivocation region  $\mathcal{R}^{(c)}$  for the model of Figure 2 with causal side information and without feedback, and they are proved in Appendix D and Appendix E.

*Theorem 3:* A single-letter characterization of the region  $\mathcal{R}^{(ci)}$  is as follows,

$$\begin{aligned}\mathcal{R}^{(ci)} &= \{(R_0, R_1, R_e) : 0 \leq R_e \leq R_1, \\ R_0 &\leq I(U; Z), \\ R_1 &\leq I(K; Y|U), \\ R_e &\leq I(K; Y|U) - I(K; Z|U)\},\end{aligned}$$

where  $p_{UKVXYZ}(u, k, v, x, y, z) = p_{Z|Y}(z|y)p_{Y|X,V}(y|x, v)p_{X|UKV}(x|u, k, v)p_{UK}(u, k)p_V(v)$ , which implies that  $(U, K) \rightarrow (X, V) \rightarrow Y \rightarrow Z$  and  $V$  is independent of  $U$  and  $K$ .

The region  $\mathcal{R}^{(ci)}$  satisfies  $\mathcal{R}^{(ci)} \subseteq \mathcal{R}^{(c)}$ .

*Remark 3:* There are some notes on Theorem 3, see the following.

- The region  $\mathcal{R}^{(ci)}$  is convex, and the proof is omitted here.
- The ranges of the random variables  $U$  and  $K$  satisfy

$$\begin{aligned}\|\mathcal{U}\| &\leq \|\mathcal{X}\|\|\mathcal{V}\| + 1, \\ \|\mathcal{K}\| &\leq (\|\mathcal{X}\|\|\mathcal{V}\| + 1)^2.\end{aligned}$$

The proof is similar to the proof of the second part of Remark 1, and it is omitted here.

- Without the secrecy parameter  $R_e$ , the region  $\mathcal{R}^{(ci)}$  is exactly the same as the capacity region for the degraded broadcast channel with causal side information [14].
- The points in  $\mathcal{R}^{(c)}$  for which  $R_e = R_1$  are of considerable interest, which imply  $H(S) = H(S|Z^N)$ .

**Definition 8: (The secrecy capacity  $C_s^{(c)}$ )** The secrecy capacity  $C_s^{(c)}$  of the model of Figure 2 with causal side information and without feedback, is denoted by

$$C_s^{(c)} = \max_{(R_0=0, R_1, R_e=R_1) \in \mathcal{R}^{(c)}} R_1. \quad (2.10)$$

Furthermore, the secrecy capacity  $C_s^{(c)}$  satisfies

$$C_s^{(c)} \geq \max(I(K; Y|U) - I(K; Z|U)). \quad (2.11)$$

*Proof of (2.11):* Substituting  $R_e = R_1$  and  $R_0 = 0$  into the region  $\mathcal{R}^{(ci)}$  in Theorem 3, we have

$$R_1 \leq I(K; Y|U),$$

$$R_1 \leq I(K; Y|U) - I(K; Z|U).$$

Note that the triple  $(R_0 = 0, R_1 = \max(I(K; Y|U) - I(K; Z|U)), R_e = R_1)$  is achievable, and therefore, the secrecy capacity  $C_s^{(c)} \geq \max(I(K; Y|U) - I(K; Z|U))$ . Thus the proof is completed. ■

*Theorem 4:* A single-letter characterization of the region  $\mathcal{R}^{(co)}$  is as follows,

$$\mathcal{R}^{(co)} = \{(R_0, R_1, R_e) : 0 \leq R_e \leq R_1,$$

$$R_0 \leq I(U; Z),$$

$$R_1 \leq I(K; Y|U),$$

$$R_e \leq I(K; Y|U) - I(K; Z|A)\},$$

where  $p_{UKAVXYZ}(u, k, a, v, x, y, z) = p_{Z|Y}(z|y)p_{Y|X,V}(y|x, v)p_{X|UAKV}(x|u, a, k, v)p_{UKA}(u, k, a)p_V(v)$ , which implies that  $(U, K, A) \rightarrow (X, V) \rightarrow Y \rightarrow Z$  and  $V$  is independent of  $U, K$  and  $A$ .

The region  $\mathcal{R}^{(co)}$  satisfies  $\mathcal{R}^{(c)} \subseteq \mathcal{R}^{(co)}$ .

*Remark 4:* There are some notes on Theorem 4, see the following.

- The region  $\mathcal{R}^{(co)}$  is convex, and the proof is omitted here.
- The ranges of the random variables  $U$  and  $K$  satisfy

$$\|\mathcal{U}\| \leq \|\mathcal{X}\| \|\mathcal{V}\| + 1,$$

$$\|\mathcal{A}\| \leq \|\mathcal{X}\| \|\mathcal{V}\|,$$

$$\|\mathcal{K}\| \leq (\|\mathcal{X}\| \|\mathcal{V}\| + 1)^2 \|\mathcal{X}\| \|\mathcal{V}\|.$$

The proof is similar to the proof of the second part of Remark 1, and it is omitted here.

- Without the secrecy parameter  $R_e$ , the region  $\mathcal{R}^{(co)}$  is exactly the same as the capacity region of the degraded broadcast channel with causal side information [14].

### III. THE MODEL OF FIGURE 2 WITH FEEDBACK

In this section, the model of Figure 2 with feedback is considered into two parts. The model of Figure 2 with noncausal side information and feedback is described in Subsection III-A, and the model of Figure 2 with causal side information and feedback is described in Subsection III-B, see the following.

#### A. The Model of Figure 2 with Noncausal Side Information and Feedback

The model of Figure 2 with noncausal side information and feedback is similar to the model in Subsection II-A, except that there is a noiseless feedback from receiver 1 to the channel encoder. The feedback  $Y^{i-1}$  (where  $2 \leq i \leq N$  and  $Y^{i-1}$  takes values in  $\mathcal{Y}^{i-1}$ ) is the previous  $i - 1$  time output of channel 1. At the  $i$ -th time,



the inputs of the channel encoder are  $S$ ,  $T$ ,  $Y^{i-1}$  and  $V^N$ , while the output is  $X_i$ , i.e., the channel encoder is a mapping

$$f_i : \mathcal{S} \times \mathcal{T} \times \mathcal{Y}^{i-1} \times \mathcal{V}^N \rightarrow \mathcal{X}_i, \quad (3.12)$$

where  $f_i(s, t, y^{i-1}, v^N) = x_i \in \mathcal{X}$ ,  $s \in \mathcal{S}$ ,  $t \in \mathcal{T}$ ,  $y^{i-1} \in \mathcal{Y}^{i-1}$  and  $v^N \in \mathcal{V}^N$ .

The following Theorem 5 and Theorem 6 provide inner and outer bounds on the capacity-equivocation region  $\mathcal{R}^{(nf)}$  for the model of Figure 2 with noncausal side information and feedback, and they are proved in Appendix F and Appendix G.

*Theorem 5:* A single-letter characterization of the region  $\mathcal{R}^{(nfi)}$  is as follows,

$$\begin{aligned} \mathcal{R}^{(nfi)} &= \{(R_0, R_1, R_e) : 0 \leq R_e \leq R_1, \\ R_0 &\leq I(U; Z) - I(U; V), \\ R_1 &\leq I(K; Y|U) - I(K; V|U), \\ R_e &\leq H(Y|Z)\}, \end{aligned}$$

where  $p_{UKVXYZ}(u, k, v, x, y, z) = p_{Z|Y}(z|y)p_{Y|X,V}(y|x, v)p_{UKXV}(u, k, x, v)$ , which implies that  $(U, K) \rightarrow (X, V) \rightarrow Y \rightarrow Z$ .

The region  $\mathcal{R}^{(nfi)}$  satisfies  $\mathcal{R}^{(nfi)} \subseteq \mathcal{R}^{(nf)}$ .

*Remark 5:* There are some notes on Theorem 5, see the following.

- The region  $\mathcal{R}^{(nfi)}$  is convex, and the proof is omitted here.
- The ranges of the random variables  $U$  and  $K$  satisfy

$$\|\mathcal{U}\| \leq \|\mathcal{X}\| \|\mathcal{V}\| + 2,$$

$$\|\mathcal{K}\| \leq (\|\mathcal{X}\| \|\mathcal{V}\| + 2)(\|\mathcal{X}\| \|\mathcal{V}\| + 1).$$

The proof is similar to the proof of the second part of Remark 1, and it is omitted here.

- Without the secrecy parameter  $R_e$ , the region  $\mathcal{R}^{(nfi)}$  is exactly the same as the achievable region for the degraded broadcast channel with noncausal side information [14].
- The points in  $\mathcal{R}^{(nf)}$  for which  $R_e = R_1$  are of considerable interest, which imply  $H(S) = H(S|Z^N)$ .

*Definition 9: (The secrecy capacity  $C_s^{(nf)}$ )* The secrecy capacity  $C_s^{(nf)}$  of the model of Figure 2 with noncausal side information and feedback, is denoted by

$$C_s^{(nf)} = \max_{(R_0=0, R_1, R_e=R_1) \in \mathcal{R}^{(nf)}} R_1. \quad (3.13)$$

Furthermore, the secrecy capacity  $C_s^{(nf)}$  satisfies

$$C_s^{(nf)} \geq \max \min \{H(Y|Z), I(K; Y|U) - I(K; V|U)\}. \quad (3.14)$$

*Proof of (3.14):* Substituting  $R_e = R_1$  and  $R_0 = 0$  into the region  $\mathcal{R}^{(nfi)}$  in Theorem 5, we have

$$R_1 \leq I(K; Y|U) - I(K; V|U),$$

$$R_1 \leq H(Y|Z).$$

Note that the triple  $(R_0 = 0, R_1 = \max \min\{H(Y|Z), I(K; Y|U) - I(K; V|U)\}, R_e = R_1)$  is achievable, and therefore, the secrecy capacity  $C_s^{(nf)} \geq \max \min\{H(Y|Z), I(K; Y|U) - I(K; V|U)\}$ . Thus the proof is completed.  $\blacksquare$

- Note that the formula  $R_e \leq I(K; Y|U) - I(K; Z|U)$  of Theorem 1 can be bounded as follows.

$$\begin{aligned} R_e &\leq I(K; Y|U) - I(K; Z|U) \\ &= H(K|U, Z) - H(K|U, Y) \\ &\stackrel{(a)}{=} H(K|U, Z) - H(K|U, Y, Z) \\ &= I(K; Y|U, Z) \leq H(Y|U, Z) \leq H(Y|Z), \end{aligned} \tag{3.15}$$

where (a) is from  $K \rightarrow (U, Y) \rightarrow Z$ .

Formula (3.15) implies that the feedback helps to enlarge the region  $\mathcal{R}^{(ni)}$  in Theorem 1.

*Theorem 6:* A single-letter characterization of the region  $\mathcal{R}^{(nfo)}$  is as follows,

$$\begin{aligned} \mathcal{R}^{(nfo)} &= \{(R_0, R_1, R_e) : 0 \leq R_e \leq R_1, \\ R_0 &\leq I(U; Z) - I(U; V), \\ R_1 &\leq I(K; Y|U, A) - I(K; V|U, A), \\ R_0 + R_1 &\leq I(U, A, K; Y) - I(U, A, K; V), \\ R_e &\leq H(Y|Z)\}, \end{aligned}$$

where  $p_{UKAVXYZ}(u, k, a, v, x, y, z) = p_{Z|Y}(z|y)p_{Y|X,V}(y|x, v)p_{UKAXV}(u, k, a, x, v)$ , which implies that  $(U, K, A) \rightarrow (X, V) \rightarrow Y \rightarrow Z$ .

The region  $\mathcal{R}^{(nfo)}$  satisfies  $\mathcal{R}^{(nf)} \subseteq \mathcal{R}^{(nfo)}$ .

*Remark 6:* There are some notes on Theorem 6, see the following.

- The region  $\mathcal{R}^{(nfo)}$  is convex, and the proof is omitted here.
- The ranges of the random variables  $U$  and  $K$  satisfy

$$\|\mathcal{U}\| \leq \|\mathcal{X}\| \|\mathcal{V}\| + 2,$$

$$\|\mathcal{A}\| \leq (\|\mathcal{X}\| \|\mathcal{V}\| + 2)(\|\mathcal{X}\| \|\mathcal{V}\| + 1),$$

$$\|\mathcal{K}\| \leq (\|\mathcal{X}\| \|\mathcal{V}\| + 2)(\|\mathcal{X}\| \|\mathcal{V}\| + 1)(\|\mathcal{X}\| \|\mathcal{V}\| + 2)^2.$$

The proof is similar to the proof of the second part of Remark 1, and it is omitted here.

- Without the secrecy parameter  $R_e$ , the region  $\mathcal{R}^{(nfo)}$  is exactly the same as the outer bound for the degraded broadcast channel with noncausal side information [14].

### B. The Model of Figure 2 with Causal Side Information and Feedback

The model of Figure 2 with causal side information and feedback is similar to the model in Subsection III-A, except that the side information  $V^N$  is known to the channel encoder in a causal manner, i.e., at the  $i$ -th time ( $1 \leq i \leq N$ ), the output of the channel encoder  $x_i = f_i(s, t, y^{i-1}, v^i)$ , where  $v^i = (v_1, v_2, \dots, v_i)$  and  $f_i$  is the mapping at time  $i$ .

The following Theorem 7 provides the capacity-equivocation region  $\mathcal{R}^{(cf)}$  for the model of Figure 2 with causal side information and feedback, and it is proved in Appendix H.

*Theorem 7:* A single-letter characterization of the capacity-equivocation region  $\mathcal{R}^{(cf)}$  is as follows,

$$\begin{aligned} \mathcal{R}^{(cf)} = \{ & (R_0, R_1, R_e) : 0 \leq R_e \leq R_1, \\ & R_0 \leq I(U; Z), \\ & R_1 \leq I(K; Y|U), \\ & R_e \leq H(Y|Z) \}, \end{aligned}$$

where  $p_{UKVXYZ}(u, k, v, x, y, z) = p_{Z|Y}(z|y)p_{Y|X,V}(y|x, v)p_{X|UKV}(x|u, k, v)p_{UK}(u, k)p_V(v)$ , which implies that  $(U, K) \rightarrow (X, V) \rightarrow Y \rightarrow Z$  and  $V$  is independent of  $U$  and  $K$ .

*Remark 7:* There are some notes on Theorem 7, see the following.

- The region  $\mathcal{R}^{(cf)}$  is convex, and the proof is omitted here.
- The ranges of the random variables  $U$  and  $K$  satisfy

$$\begin{aligned} \|\mathcal{U}\| & \leq \|\mathcal{X}\|\|\mathcal{V}\| + 1, \\ \|\mathcal{K}\| & \leq (\|\mathcal{X}\|\|\mathcal{V}\| + 1)\|\mathcal{X}\|\|\mathcal{V}\|. \end{aligned}$$

The proof is similar to the proof of the second part of Remark 1, and it is omitted here.

- Without the secrecy parameter  $R_e$ , the region  $\mathcal{R}^{(cf)}$  is exactly the same as the capacity region for the degraded broadcast channel with causal side information [14], i.e., **the noiseless feedback can not increase the capacity of the degraded broadcast channel with causal side information.**
- The points in  $\mathcal{R}^{(cf)}$  for which  $R_e = R_1$  are of considerable interest, which imply  $H(S) = H(S|Z^N)$ .

*Definition 10:* (**The secrecy capacity**  $C_s^{(cf)}$ ) The secrecy capacity  $C_s^{(cf)}$  of the model of Figure 2 with causal side information and feedback, is denoted by

$$C_s^{(cf)} = \max_{(R_0=0, R_1, R_e=R_1) \in \mathcal{R}^{(cf)}} R_1. \quad (3.16)$$

Furthermore, the secrecy capacity  $C_s^{(cf)}$  satisfies

$$C_s^{(cf)} = \max \min \{I(K; Y|U), H(Y|Z)\}. \quad (3.17)$$

*Proof of (3.17):* Substituting  $R_e = R_1$  and  $R_0 = 0$  into the region  $\mathcal{R}^{(cf)}$  in Theorem 7, we have

$$R_1 \leq I(K; Y|U), \quad (3.18)$$

$$R_1 \leq H(Y|Z). \quad (3.19)$$

By using (3.16), (3.18) and (3.19), (3.17) is proved.  $\blacksquare$

- By using the same formula as (3.15), it is easy to see that the feedback helps to enlarge the region  $\mathcal{R}^{(ci)}$  in Theorem 3.

#### IV. A BINARY EXAMPLE FOR THE MODEL OF FIGURE 2 WITH CAUSAL SIDE INFORMATION AND NOISELESS FEEDBACK

In this section, we calculate the secrecy capacity of a special case of the model of Figure 2 with causal side information and noiseless feedback.

Suppose that the channel state information  $V^N$  is available at the channel encoder in a casual manner, and the random variable  $V$  is uniformly distributed over  $\{0, 1\}$ , i.e.,  $p_V(0) = p_V(1) = \frac{1}{2}$ . Meanwhile, the random variables  $X$ ,  $Y$  and  $Z$  take values in  $\{0, 1\}$ , and channel 2 is a BSC (binary symmetric channel) with crossover probability  $q$ . The transition probability of channel 1 is defined as follows:

When  $v = 0$ ,

$$p_{Y|X,V}(y|x, v = 0) = \begin{cases} 1 - p, & \text{if } y = x, \\ p, & \text{otherwise.} \end{cases} \quad (4.1)$$

When  $v = 1$ ,

$$p_{Y|X,V}(y|x, v = 1) = \begin{cases} p, & \text{if } y = x, \\ 1 - p, & \text{otherwise.} \end{cases} \quad (4.2)$$

From Remark 7 we know that the secrecy capacity for the causal case is

$$C_s^{(cf)} = \max \min \{I(K; Y|U), H(Y|Z)\}. \quad (4.3)$$

Since  $R_0 = 0$ , we have  $U = \text{const}$ , and therefore,  $I(K; Y|U) = I(K; Y)$ . The remainder of this example is to calculate the characters  $\max I(K; Y)$  and  $\max H(Y|Z)$ .

Let  $K$  take values in  $\{0, 1\}$ . The probability of  $K$  is defined as follows.  $p_K(0) = \alpha$  and  $p_K(1) = 1 - \alpha$ .

In addition, define the conditional probability mass function  $p_{X|K,V}$  as follows.

$$p_{X|K,V}(0|0, 0) = \beta_1, p_{X|K,V}(1|0, 0) = 1 - \beta_1, p_{X|K,V}(0|0, 1) = \beta_2, p_{X|K,V}(1|0, 1) = 1 - \beta_2,$$

$$p_{X|K,V}(0|1, 0) = \beta_3, p_{X|K,V}(1|1, 0) = 1 - \beta_3, p_{X|K,V}(0|1, 1) = \beta_4, p_{X|K,V}(1|1, 1) = 1 - \beta_4.$$

The character  $I(K; Y)$  depends on the joint probability mass functions  $p_{KY}$ , and we have

$$\begin{aligned} p_{KY}(k, y) &= \sum_{x,v} p_{K Y X V}(k, y, x, v) \\ &= \sum_{x,v} p_{Y|X V}(y|x, v) p_{X|K,V}(x|k, v) p_K(k) p_V(v). \end{aligned} \quad (4.4)$$

Now we calculate  $\max I(K; Y)$  and  $\max H(Y|Z)$ , respectively.

- (Calculation of  $\max I(K; Y)$ ) Since

$$I(K; Y) = \sum_k \sum_y p_{KY}(k, y) \log \frac{p_{KY}(k, y)}{p_K(k)p_Y(y)}, \quad (4.5)$$

and

$$p_{KY}(0, 0) = \frac{\alpha}{2}[1 - (\beta_1 - \beta_2)(1 - 2p)], \quad (4.6)$$

$$p_{KY}(0, 1) = \frac{\alpha}{2}[1 + (\beta_1 - \beta_2)(1 - 2p)], \quad (4.7)$$

$$p_{KY}(1, 0) = \frac{\alpha}{2}[1 - (\beta_3 - \beta_4)(1 - 2p)], \quad (4.8)$$

$$p_{KY}(1, 1) = \frac{\alpha}{2}[1 + (\beta_3 - \beta_4)(1 - 2p)]. \quad (4.9)$$

Define  $\beta_1 - \beta_2 = a$  and  $\beta_3 - \beta_4 = b$ , then we have

$$I(K; Y) \leq h(1 - \alpha - p + 2p\alpha) - h(p) \leq 1 - h(p), \quad (4.10)$$

where  $h(x) = -x \log x - (1 - x) \log(1 - x)$ ,  $0 \leq x \leq 1$ , and the “=” is achieved if  $a = 1$ ,  $b = -1$  and  $\alpha = \frac{1}{2}$ .

- The character  $H(Y|Z)$  is calculated as follows.

$$\begin{aligned} H(Y|Z) &= H(Z|Y) + H(Y) - H(Z) \\ &\leq H(Z|Y) = h(q), \end{aligned} \quad (4.11)$$

and therefore,  $\max H(Y|Z) = h(q)$ . Note that  $a = 1$ ,  $b = -1$  and  $\alpha = \frac{1}{2}$  are also the distributions for  $H(Y|Z) = h(q)$ .

Thus, the secrecy capacity of this special case is

$$C_s^{(cf)} = \min\{1 - h(p), h(q)\}. \quad (4.12)$$

## V. CONCLUSION

In this paper, we investigate the model of degraded broadcast channel with side information, confidential messages and with or without noiseless feedback. This work is from Steinberg’s work on the degraded broadcast channel with causal and noncausal side information, and Csiszár and Körner’s work on broadcast channel with confidential messages. For the non-feedback model, inner and outer bounds on the capacity-equivocation regions are provided for both causal and noncausal manners. Superposition coding, Gel’fand-Pinsker’s binning and Wyner’s random binning technique are used in the corresponding achievability proofs.

For the feedback model, inner and outer bounds on the capacity-equivocation region are provided for the noncausal case, and the capacity-equivocation region is determined for the causal case. In the corresponding achievability proofs, the noiseless feedback is used as a secret key shared by receiver 1 and transmitter, and therefore, the coding schemes for the achievability proofs are a combination of superposition coding, Gel’fand-Pinsker’s binning, block Markov coding and Ahlswede-Cai’s secret key on the feedback system.

Finally, we give an example on calculating the secrecy capacity of the binary degraded broadcast channel with causal side information, confidential messages and noiseless feedback.

## ACKNOWLEDGEMENT

The authors would like to thank Professor Yossef Steinberg for his valuable suggestions to improve this paper.

## APPENDIX A

## PROOF OF THEOREM 1

In this section, we will show that any triple  $(R_0, R_1, R_e) \in \mathcal{R}^{ni}$  is achievable. Superposition coding, Gel'fand-Pinsker's binning and Wyner's random binning technique are used in the construction of the code-books.

Now the remainder of this section is organized as follows. The code construction is in Subsection A-A. The proof of achievability is given in Subsection A-B.

## A. Code Construction

Since  $R_e \leq I(K; Y|U) - I(K; Z|U)$  and  $R_e \leq R_1 \leq I(K; Y|U) - I(K; V|U)$ , it is sufficient to show that the triple  $(R_0, R_1, R_e = I(K; Y|U) - \max(I(K; Z|U), I(K; V|U)))$  is achievable, and note that this implies that  $R_1 \geq R_e = I(K; Y|U) - \max(I(K; Z|U), I(K; V|U))$ .

Given a triple  $(R_0, R_1, R_e)$ , choose a joint probability mass function  $p_{U,K,V,X,Y,Z}(u, k, v, x, y, z)$  such that

$$0 \leq R_e \leq R_1,$$

$$R_0 \leq I(U; Z) - I(U; V),$$

$$R_1 \leq I(K; Y|U) - I(K; V|U),$$

$$R_e = I(K; Y|U) - \max(I(K; Z|U), I(K; V|U)).$$

The confidential message set  $\mathcal{S}$  and the common message set  $\mathcal{T}$  satisfy the following conditions:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \|\mathcal{S}\| = R_1 = I(K; Y|U) - I(K; V|U) - \gamma_1, \quad (13)$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \|\mathcal{T}\| = R_0 = I(U; Z) - I(U; V) - \gamma, \quad (14)$$

where  $\gamma$  and  $\gamma_1$  are fixed positive real numbers and

$$0 \leq \gamma_1 \leq^{(a)} \max(I(K; Z|U), I(K; V|U)) - I(K; V|U). \quad (15)$$

Note that (a) is from  $R_1 \geq R_e = I(K; Y|U) - \max(I(K; Z|U), I(K; V|U))$  and (13). Let  $\mathcal{S} = \{1, 2, \dots, 2^{NR_1}\}$  and  $\mathcal{T} = \{1, 2, \dots, 2^{NR_0}\}$ .

Code-book generation:

- **(Construction of  $U^N$ )** Gel'fand-Pinsker's binning technique is used in the construction of  $U^N$ , see Figure 3. Generate  $2^{N(I(U;Z)-\epsilon_{1,N})}$  ( $\epsilon_{1,N} \rightarrow 0$  as  $N \rightarrow \infty$ ) i.i.d. sequences  $u^N$ , according to the probability mass function  $p_U(u)$ . Distribute these sequences at random into  $2^{NR_0} = 2^{N(I(U;Z)-I(U;V)-\gamma)}$  bins such that each bin contains  $2^{N(I(U;V)+\gamma-\epsilon_{1,N})}$  sequences. Index each bin by  $i \in \{1, 2, \dots, 2^{NR_0}\}$ .

For a given common message  $t$  ( $t \in \mathcal{T}$ ) and side information  $v^N$ , try to find a sequence in bin  $t$   $\{u^N(t, 1), u^N(t, 2), \dots, u^N(t, 2^{N(I(U;V)+\gamma-\epsilon_{1,N})})\}$  that is jointly typical with  $v^N$ , say  $u^N(t, i^*)$ , i.e.,  $(u^N(t, i^*), v^N) \in T_{UV}^N(\epsilon_1)$ . If multiple such sequences in bin  $t$  exist, choose the one with the smallest  $i^*$ . If no such  $i^*$  exists, then declare an encoding error.

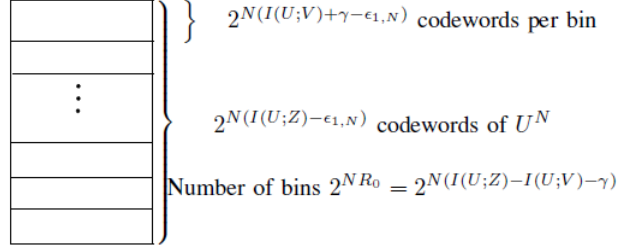


Fig. 3: Code-book construction for  $U^N$  in Theorem 1

- **(Construction of  $K^N$ )** Classical superposition coding and double binning technique [3] are used in the construction of  $K^N$ , see Figure 4.

For the transmitted sequence  $u^N(t, i^*)$ , generate  $2^{N(I(K;Y|U)-\epsilon_{2,N})}$  ( $\epsilon_{2,N} \rightarrow 0$  as  $N \rightarrow \infty$ ) i.i.d. sequences  $k^N$ , according to the probability mass function  $p_{K|U}(k_i|u_i(t, i^*))$ . Distribute these sequences at random into  $2^{NR_1} = 2^{N(I(K;Y|U)-I(K;V|U)-\gamma_1)}$  bins such that each bin contains  $2^{N(I(K;V|U)+\gamma_1-\epsilon_{2,N})}$  sequences. Index each bin by  $i \in \{1, 2, \dots, 2^{NR_1}\}$ . Then place the  $2^{N(I(K;V|U)+\gamma_1-\epsilon_{2,N})}$  sequences in every bin randomly into  $2^{N(\max(I(K;V|U), I(K;Z|U))-I(K;Z|U)+\epsilon_{3,N})}$  ( $\epsilon_{3,N} \rightarrow 0$  as  $N \rightarrow \infty$ ) subbins such that every subbin contains  $2^{N(I(K;V|U)+\gamma_1-\epsilon_{2,N}-\max(I(K;V|U), I(K;Z|U))+I(K;Z|U)-\epsilon_{3,N})}$  sequences. Let  $J$  be the random variable to represent the index of the subbin. Index each subbin by  $j \in \{1, 2, \dots, 2^{N(\max(I(K;V|U), I(K;Z|U))-I(K;Z|U)+\epsilon_{3,N})}\}$ , i.e.,

$$\log \|\mathcal{J}\| = N(\max(I(K;V|U), I(K;Z|U)) - I(K;Z|U) + \epsilon_{3,N}). \quad (16)$$

Here note that the number of the sequences in every subbin is upper bounded as follows.

$$\begin{aligned} & I(K;V|U) + \gamma_1 - \epsilon_{2,N} - \max(I(K;V|U), I(K;Z|U)) + I(K;Z|U) - \epsilon_{3,N} \\ & \stackrel{(a)}{\leq} I(K;Z|U) - \epsilon_{2,N} - \epsilon_{3,N}, \end{aligned} \quad (17)$$

where (a) is from (15). This implies that

$$\lim_{N \rightarrow \infty} H(K^N|U^N, S, J, Z^N) = 0. \quad (18)$$

For a given confidential message  $s$  ( $s \in \mathcal{S}$ ) and side information  $v^N$ , try to find a sequence  $k^N(u^N(t, i^*))$  in bin  $s$  such that  $(k^N(u^N(t, i^*)), v^N) \in T_{KV|U}^N(\epsilon_2)$ . If multiple such sequences in bin  $s$  exist, choose the one with the smallest index in the bin. If no such sequence exists, declare an encoding error.

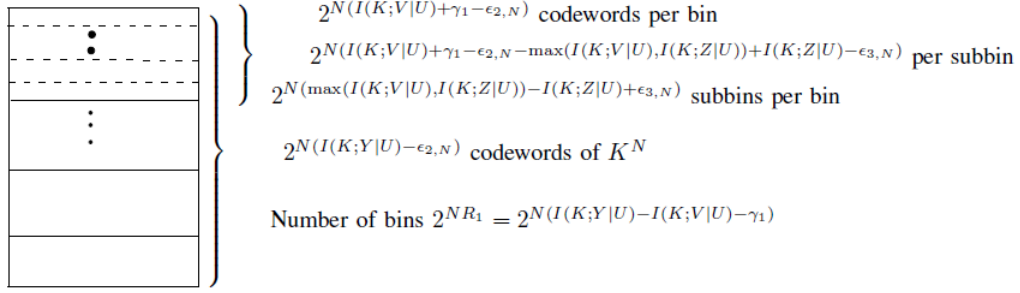


Fig. 4: Code-book construction for  $K^N$  in Theorem 1

- **(Construction of  $X^N$ )** The  $x^N$  is generated according to a new discrete memoryless channel (DMC) with inputs  $k^N, u^N, v^N$ , and output  $x^N$ . The transition probability of this new DMC is  $p_{X|U,K,V}(x|u, k, v)$ , which is obtained from the joint probability mass function  $p_{U,K,V,X,Y,Z}(u, k, v, x, y, z)$ . The probability  $p_{X^N|U^N, K^N, V^N}(x^N|u^N, k^N, v^N)$  is calculated as follows.

$$p_{X^N|U^N, K^N, V^N}(x^N|u^N, k^N, v^N) = \prod_{i=1}^N p_{X|U,K,V}(x_i|u_i, k_i, v_i). \quad (19)$$

Decoding:

Receiver 2: Given a vector  $z^N \in \mathcal{Z}^N$ , try to find a sequence  $u^N(\hat{t}, \hat{i})$  such that  $(u^N(\hat{t}, \hat{i}), z^N) \in T_{UZ}^N(\epsilon_3)$ . If there exist sequences with the same  $\hat{t}$ , put out the corresponding  $\hat{t}$ . Otherwise, i.e., if no such sequence exists or multiple sequences have different message indices, declare a decoding error.

Receiver 1: Given a vector  $y^N \in \mathcal{Y}^N$ , try to find a sequence  $u^N$  such that  $(u^N, y^N) \in T_{UY}^N(\epsilon_4)$ . If such a sequence does not exist, or there are more than one such sequence, declare a decoding error. Denote the corresponding sequence by  $u^N(\hat{t}, \hat{i})$ , put out the corresponding index  $\hat{t}$ .

After decoding  $u^N(\hat{t}, \hat{i})$  and  $\hat{t}$ , try to find a sequence  $k^N(u^N(\hat{t}, \hat{i}))$  such that  $(k^N(u^N(\hat{t}, \hat{i})), y^N) \in T_{KY|U}^N(\epsilon_5)$ . If there exist sequences with the same index of the bin  $\hat{s}$ , put out the corresponding  $\hat{s}$ . Otherwise, declare a decoding error.

### B. Proof of Achievability

By using the above definitions, it is easy to verify that  $\lim_{N \rightarrow \infty} \frac{\log \|T\|}{N} = R_0$  and  $\lim_{N \rightarrow \infty} \frac{\log \|S\|}{N} = R_1$ .

Then, note that the above encoding and decoding scheme is similar to the one used in [14]. Hence, by similar arguments as in [14], it is easy to show that  $P_{e1} \leq \epsilon$  and  $P_{e2} \leq \epsilon$ , and the proof is omitted here. It remains to show that  $\lim_{N \rightarrow \infty} \Delta \geq R_e$ , see the following.



$$\begin{aligned}
\lim_{N \rightarrow \infty} \Delta &= \lim_{N \rightarrow \infty} \frac{1}{N} H(S|Z^N) \\
&\geq \lim_{N \rightarrow \infty} \frac{1}{N} H(S|Z^N, U^N) \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} (H(S, Z^N, U^N) - H(Z^N, U^N)) \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} (H(S, Z^N, U^N, J, K^N) - H(J, K^N|Z^N, U^N, S) - H(Z^N, U^N)) \\
&\stackrel{(a)}{=} \lim_{N \rightarrow \infty} \frac{1}{N} (H(Z^N|U^N, K^N) + H(U^N, J, K^N, S) - H(J, K^N|Z^N, U^N, S) - H(Z^N, U^N)) \\
&\stackrel{(b)}{=} \lim_{N \rightarrow \infty} \frac{1}{N} (H(Z^N|U^N, K^N) + H(U^N, K^N) - H(J, K^N|Z^N, U^N, S) - H(Z^N, U^N)) \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} (H(K^N|U^N) - H(J, K^N|Z^N, U^N, S) - I(Z^N; K^N|U^N)) \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} (H(K^N|U^N) - H(J|Z^N, U^N, S) - H(K^N|Z^N, U^N, S, J) - I(Z^N; K^N|U^N)) \\
&\stackrel{(c)}{\geq} \lim_{N \rightarrow \infty} \frac{1}{N} (H(K^N|U^N) - \log \|\mathcal{J}\| - H(K^N|Z^N, U^N, S, J) - I(Z^N; K^N|U^N)) \\
&\geq \lim_{N \rightarrow \infty} \frac{1}{N} (H(K^N|U^N) - H(K^N|U^N, Y^N) - \log \|\mathcal{J}\| - H(K^N|Z^N, U^N, S, J) - I(Z^N; K^N|U^N)) \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} (I(Y^N; K^N|U^N) - \log \|\mathcal{J}\| - H(K^N|Z^N, U^N, S, J) - I(Z^N; K^N|U^N)) \\
&\stackrel{(d)}{=} \lim_{N \rightarrow \infty} \frac{1}{N} (NI(Y; K|U) - \log \|\mathcal{J}\| - H(K^N|Z^N, U^N, S, J) - NI(Z; K|U)) \\
&\stackrel{(e)}{=} \lim_{N \rightarrow \infty} \frac{1}{N} (NI(Y; K|U) - N \max(I(K; V|U), I(K; Z|U)) + NI(K; Z|U) - N\epsilon_{3,N} - NI(Z; K|U)) \\
&\stackrel{(f)}{=} I(Y; K|U) - \max(I(K; V|U), I(K; Z|U)) = R_e, \tag{20}
\end{aligned}$$

where (a) is from  $(S, J) \rightarrow (U^N, K^N) \rightarrow Z^N$ , (b) is from  $H(J, S|U^N, K^N) = 0$ , (c) is from  $H(J|Z^N, U^N, S) \leq H(J) \leq \log \|\mathcal{J}\|$ , (d) is from that  $V^N, U^N, K^N$  and  $X^N$  are i.i.d. generated random vectors, and the channels are discrete memoryless, (e) is from (16) and (18), and (f) is from  $\epsilon_{3,N} \rightarrow 0$  as  $N \rightarrow \infty$ .

Thus,  $\lim_{N \rightarrow \infty} \Delta \geq R_e$  is proved, and the proof of Theorem 1 is completed.

## APPENDIX B

### PROOF OF THEOREM 2

In this section, we prove Theorem 2: all the achievable  $(R_0, R_1, R_e)$  triples are contained in the set  $\mathcal{R}^{(no)}$ . Suppose  $(R_0, R_1, R_e)$  is achievable, i.e., for any given  $\epsilon > 0$ , there exists a channel encoder-decoder  $(N, \Delta, P_{e1}, P_{e2})$  such that

$$\lim_{N \rightarrow \infty} \frac{\log \|\mathcal{T}\|}{N} = R_0, \lim_{N \rightarrow \infty} \frac{\log \|\mathcal{S}\|}{N} = R_1, \lim_{N \rightarrow \infty} \Delta \geq R_e, P_{e1} \leq \epsilon, P_{e2} \leq \epsilon.$$

Then we will show the existence of random variables  $(U, K, A) \rightarrow (X, V) \rightarrow Y \rightarrow Z$  such that

$$0 \leq R_e \leq R_1, \tag{21}$$

$$R_0 \leq I(U; Z) - I(U; V), \tag{22}$$

$$R_1 \leq I(K; Y|U, A) - I(K; V|U, A), \quad (23)$$

$$R_0 + R_1 \leq I(U, K, A; Y) - I(U, K, A; V), \quad (24)$$

$$R_e \leq I(K; Y|U, A) - I(K; V|U, A) - I(K; Z|U) + I(K; V|U). \quad (25)$$

Since  $S$  and  $T$  are independent and uniformly distributed over  $\mathcal{S}$  and  $\mathcal{T}$ , we have  $H(S) = \log \|\mathcal{S}\|$  and  $H(T) = \log \|\mathcal{T}\|$ . The formulas (22), (23), (24) and (25) are proved by Lemma 1, see the following.

*Lemma 1:* The random vectors  $Y^N$ ,  $Z^N$  and the random variables  $S$ ,  $T$ ,  $V$ ,  $U$ ,  $K$ ,  $A$ ,  $Y$ ,  $Z$  of Theorem 2, satisfy:

$$\frac{1}{N}H(T) \leq I(U; Z) - I(U; V) + \frac{1}{N}\delta(P_{e2}), \quad (26)$$

$$\frac{1}{N}H(S) = \frac{1}{N}H(S|T) \leq I(K; Y|U, A) - I(K; V|U, A) + \frac{1}{N}\delta(P_{e1}), \quad (27)$$

$$\frac{1}{N}H(S, T) \leq I(U, K, A; Y) - I(U, K, A; V) + \frac{1}{N}\delta(P_{e1}), \quad (28)$$

$$\begin{aligned} \frac{1}{N}H(S|Z^N) &\leq I(K; Y|U, A) - I(K; V|U, A) - I(K; Z|U) + I(K; V|U) \\ &\quad + \frac{1}{N}\delta(P_{e1}) + \frac{1}{N}\delta(P_{e2}), \end{aligned} \quad (29)$$

where  $\delta(P_{e1}) = h(P_{e1}) + P_{e1} \log(|\mathcal{S} \times \mathcal{T}| - 1)$  and  $\delta(P_{e2}) = h(P_{e2}) + P_{e2} \log(|\mathcal{T}| - 1)$ . Note that  $h(P_{e1}) = -P_{e1} \log P_{e1} - (1 - P_{e1}) \log(1 - P_{e1})$  and  $h(P_{e2}) = -P_{e2} \log P_{e2} - (1 - P_{e2}) \log(1 - P_{e2})$ .

Substituting  $H(S) = \log \|\mathcal{S}\|$ ,  $H(T) = \log \|\mathcal{T}\|$ ,  $H(S, T) = H(S) + H(T)$  and (2.4) into (26), (27), (28) and (29), and using the fact that  $\epsilon \rightarrow 0$ , the formulas (22), (23), (24) and (25) are obtained. The formula (21) is from

$$R_e \leq \lim_{N \rightarrow \infty} \Delta = \lim_{N \rightarrow \infty} \frac{1}{N}H(S|Z^N) \leq \lim_{N \rightarrow \infty} \frac{1}{N}H(S) = R_1.$$

It remains to prove Lemma 1, see the following.

*Proof of Lemma 1:* The formula (26) follows from (30), (34) and (44). The formula (27) is from (31), (36) and (48). The formula (28) is from (32), (38) and (49). The formula (29) is proved by (33), (36), (39), (48) and (50).

<Part i> We begin with the left parts of the inequalities (26), (27), (28) and (29), see the following.

Since  $T \rightarrow Y^N \rightarrow Z^N$  is a Markov chain, for the common message  $T$ , we have

$$\begin{aligned} \frac{1}{N}H(T) &= \frac{1}{N}H(T|Z^N) + \frac{1}{N}I(Z^N; T) \\ &\stackrel{(a)}{\leq} \frac{1}{N}\delta(P_{e2}) + \frac{1}{N}I(Z^N; T). \end{aligned} \quad (30)$$

For the confidential message  $S$ , we have

$$\begin{aligned} \frac{1}{N}H(S) &= \frac{1}{N}H(S|T) = \frac{1}{N}I(S; Y^N|T) + \frac{1}{N}H(S|Y^N, T) \\ &\leq \frac{1}{N}I(S; Y^N|T) + \frac{1}{N}H(S, T|Y^N) \\ &\stackrel{(b)}{\leq} \frac{1}{N}I(S; Y^N|T) + \frac{1}{N}\delta(P_{e1}). \end{aligned} \quad (31)$$

For  $S$  and  $T$ , we have

$$\begin{aligned} \frac{1}{N}(H(S) + H(T)) &= \frac{1}{N}H(S, T) = \frac{1}{N}I(S, T; Y^N) + \frac{1}{N}H(S, T|Y^N) \\ &\stackrel{(c)}{\leq} \frac{1}{N}I(S, T; Y^N) + \frac{1}{N}\delta(P_{e1}). \end{aligned} \quad (32)$$

For the equivocation to the receiver 2, we have

$$\begin{aligned} \frac{1}{N}H(S|Z^N) &= \frac{1}{N}H(S|Z^N, T) + \frac{1}{N}I(S; T|Z^N) \\ &= \frac{1}{N}H(S|T) - \frac{1}{N}I(S; Z^N|T) + \frac{1}{N}I(S; T|Z^N) \\ &= \frac{1}{N}I(S; Y^N|T) + \frac{1}{N}H(S|Y^N, T) - \frac{1}{N}I(S; Z^N|T) + \frac{1}{N}I(S; T|Z^N) \\ &\leq \frac{1}{N}I(S; Y^N|T) - \frac{1}{N}I(S; Z^N|T) + \frac{1}{N}H(S, T|Y^N) + \frac{1}{N}H(T|Z^N) \\ &\stackrel{(d)}{\leq} \frac{1}{N}I(S; Y^N|T) - \frac{1}{N}I(S; Z^N|T) + \frac{1}{N}\delta(P_{e1}) + \frac{1}{N}\delta(P_{e2}). \end{aligned} \quad (33)$$

Note that (a), (b), (c) and (d) follow from Fano's inequality.

<Part ii> By using chain rule, the character  $I(Z^N; T)$  in formula (30) can be bounded as follows,

$$\begin{aligned} \frac{1}{N}I(Z^N; T) &= \frac{1}{N} \sum_{i=1}^N I(Z_i; T|Z^{i-1}) \\ &\stackrel{(1)}{=} \frac{1}{N} \sum_{i=1}^N (I(Z_i; T|Z^{i-1}) - I(V_i; T|V_{i+1}^N)) \\ &= \frac{1}{N} \sum_{i=1}^N (I(Z_i; T, V_{i+1}^N|Z^{i-1}) - I(Z_i; V_{i+1}^N|T, Z^{i-1}) - I(V_i; T, Z^{i-1}|V_{i+1}^N) + I(V_i; Z^{i-1}|T, V_{i+1}^N)) \\ &\stackrel{(2)}{=} \frac{1}{N} \sum_{i=1}^N (I(Z_i; T, V_{i+1}^N|Z^{i-1}) - I(V_i; T, Z^{i-1}|V_{i+1}^N)) \\ &\stackrel{(3)}{=} \frac{1}{N} \sum_{i=1}^N (H(Z_i|Z^{i-1}) - H(Z_i|Z^{i-1}, T, V_{i+1}^N) - H(V_i) + H(V_i|Z^{i-1}, T, V_{i+1}^N)) \\ &\leq \frac{1}{N} \sum_{i=1}^N (H(Z_i) - H(Z_i|Z^{i-1}, T, V_{i+1}^N) - H(V_i) + H(V_i|Z^{i-1}, T, V_{i+1}^N)), \end{aligned} \quad (34)$$

where formula (1) follows from that  $V_i(1 \leq i \leq N)$  are i.i.d. random variables and they are independent of  $T$ ,

formula (2) follows from that

$$\sum_{i=1}^N I(Z_i; V_{i+1}^N|T, Z^{i-1}) = \sum_{i=1}^N I(V_i; Z^{i-1}|T, V_{i+1}^N), \quad (35)$$

and formula (3) follows from that  $V_i(1 \leq i \leq N)$  are i.i.d. random variables.

<Part iii> Using chain rule, the character  $I(S; Y^N|T)$  in formulas (31) and (33) can be rewritten as follows,

$$\begin{aligned}
\frac{1}{N}I(Y^N; S|T) &= \frac{1}{N} \sum_{i=1}^N I(Y_i; S|Y^{i-1}, T) \\
&\stackrel{(a)}{=} \frac{1}{N} \sum_{i=1}^N (I(Y_i; S|Y^{i-1}, T) - I(V_i; S|V_{i+1}^N, T)) \\
&= \frac{1}{N} \sum_{i=1}^N (I(Y_i; S, V_{i+1}^N|Y^{i-1}, T) - I(Y_i; V_{i+1}^N|T, Y^{i-1}, S) - I(V_i; S, Y^{i-1}|V_{i+1}^N, T) + \\
&\quad I(V_i; Y^{i-1}|T, S, V_{i+1}^N)) \\
&\stackrel{(b)}{=} \frac{1}{N} \sum_{i=1}^N (I(Y_i; S, V_{i+1}^N|Y^{i-1}, T) - I(V_i; S, Y^{i-1}|V_{i+1}^N, T)) \\
&= \frac{1}{N} \sum_{i=1}^N (H(S, V_{i+1}^N|Y^{i-1}, T) - H(S, V_{i+1}^N|Y^{i-1}, T, Y_i) - H(S, Y^{i-1}|V_{i+1}^N, T) + \\
&\quad H(S, Y^{i-1}|V_{i+1}^N, T, V_i)) \\
&= \frac{1}{N} \sum_{i=1}^N (H(V_{i+1}^N|Y^{i-1}, T) + H(S|Y^{i-1}, T, V_{i+1}^N) - H(V_{i+1}^N|Y^{i-1}, T, Y_i) - \\
&\quad H(S|Y^{i-1}, T, Y_i, V_{i+1}^N) - H(Y^{i-1}|V_{i+1}^N, T) - H(S|V_{i+1}^N, T, Y^{i-1}) + \\
&\quad H(Y^{i-1}|V_{i+1}^N, T, V_i) + H(S|V_{i+1}^N, T, V_i, Y^{i-1})) \\
&= \frac{1}{N} \sum_{i=1}^N (I(V_{i+1}^N; Y_i|Y^{i-1}, T) - I(Y^{i-1}; V_i|V_{i+1}^N, T) + \\
&\quad I(S; Y_i|Y^{i-1}, T, V_{i+1}^N) - I(S; V_i|Y^{i-1}, V_{i+1}^N, T)) \\
&\stackrel{(c)}{=} \frac{1}{N} \sum_{i=1}^N (I(S; Y_i|Y^{i-1}, T, V_{i+1}^N) - I(S; V_i|Y^{i-1}, V_{i+1}^N, T)) \\
&\stackrel{(d)}{=} \frac{1}{N} \sum_{i=1}^N (I(S; Y_i|Y^{i-1}, T, V_{i+1}^N, Z^{i-1}) - I(S; V_i|Y^{i-1}, V_{i+1}^N, T, Z^{i-1})), \tag{36}
\end{aligned}$$

where formula (a) follows from  $V_i (1 \leq i \leq N)$  are i.i.d. random variables and they are independent of  $T, S$ , formula (b) follows from

$$\sum_{i=1}^N I(Y_i; V_{i+1}^N|T, Y^{i-1}, S) = \sum_{i=1}^N I(V_i; Y^{i-1}|T, S, V_{i+1}^N), \tag{37}$$

formula (c) follows from (35), and formula (d) is from the Markov chains  $Y_i \rightarrow (Y^{i-1}, T, V_{i+1}^N) \rightarrow Z^{i-1}$ ,  $Y_i \rightarrow (Y^{i-1}, T, V_{i+1}^N, S) \rightarrow Z^{i-1}$ ,  $V_i \rightarrow (Y^{i-1}, T, V_{i+1}^N) \rightarrow Z^{i-1}$  and  $V_i \rightarrow (Y^{i-1}, T, V_{i+1}^N, S) \rightarrow Z^{i-1}$ .

<Part iv> Similar to (34), the character  $I(S, T; Y^N)$  in formula (32) can be rewritten as follows,

$$\frac{1}{N}I(S, T; Y^N) \leq \frac{1}{N} \sum_{i=1}^N (H(Y_i) - H(Y_i|Y^{i-1}, T, S, V_{i+1}^N) - H(V_i) + H(V_i|Y^{i-1}, T, S, V_{i+1}^N)). \tag{38}$$

<Part v> Similar to (36), the character  $I(S; Z^N|T)$  in formula (33) can be rewritten as follows,

$$\begin{aligned}
\frac{1}{N}I(S; Z^N|T) &= \frac{1}{N} \sum_{i=1}^N I(Z_i; S|Z^{i-1}, T) \\
&= \frac{1}{N} \sum_{i=1}^N (I(Z_i; S|Z^{i-1}, T) - I(V_i; S|V_{i+1}^N, T)) \\
&= \frac{1}{N} \sum_{i=1}^N (I(Z_i; S, V_{i+1}^N|Z^{i-1}, T) - I(Z_i; V_{i+1}^N|T, Z^{i-1}, S) - I(V_i; S, Z^{i-1}|V_{i+1}^N, T) + \\
&\quad I(V_i; Z^{i-1}|T, S, V_{i+1}^N)) \\
&= \frac{1}{N} \sum_{i=1}^N (I(Z_i; S, V_{i+1}^N|Z^{i-1}, T) - I(V_i; S, Z^{i-1}|V_{i+1}^N, T)) \\
&= \frac{1}{N} \sum_{i=1}^N (I(V_{i+1}^N; Z_i|Z^{i-1}, T) - I(Z^{i-1}; V_i|V_{i+1}^N, T) + \\
&\quad I(S; Z_i|Z^{i-1}, T, V_{i+1}^N) - I(S; V_i|Z^{i-1}, V_{i+1}^N, T)) \\
&= \frac{1}{N} \sum_{i=1}^N (I(S; Z_i|Z^{i-1}, T, V_{i+1}^N) - I(S; V_i|Z^{i-1}, V_{i+1}^N, T)). \tag{39}
\end{aligned}$$

<Part vi> (single letter) To complete the proof, we introduce a random variable  $J$ , which is independent of  $S$ ,  $T$ ,  $X^N$ ,  $V^N$ ,  $Y^N$  and  $Z^N$ . Furthermore,  $J$  is uniformly distributed over  $\{1, 2, \dots, N\}$ . Define

$$U = (T, Z^{J-1}, V_{J+1}^N, J), \tag{40}$$

$$K = S, \tag{41}$$

$$A = (Y^{J-1}, J), \tag{42}$$

$$X = X_J, Y = Y_J, Z = Z_J, V = V_J. \tag{43}$$

<Part vii> Then (34) can be rewritten as

$$\begin{aligned}
\frac{1}{N}I(T; Z^N) &\leq \frac{1}{N} \sum_{i=1}^N (H(Z_i) - H(Z_i|Z^{i-1}, T, V_{i+1}^N) - H(V_i) + H(V_i|Z^{i-1}, T, V_{i+1}^N)) \\
&= \frac{1}{N} \sum_{i=1}^N (H(Z_i|J=i) - H(Z_i|Z^{i-1}, T, V_{i+1}^N, J=i) - H(V_i|J=i) + H(V_i|Z^{i-1}, T, V_{i+1}^N, J=i)) \\
&= H(Z_J|J) - H(Z_J|Z^{J-1}, T, V_{J+1}^N, J) - H(V_J|J) + H(V_J|Z^{J-1}, T, V_{J+1}^N, J) \\
&\stackrel{(a)}{=} H(Z_J|J) - H(Z_J|Z^{J-1}, T, V_{J+1}^N, J) - H(V_J) + H(V_J|Z^{J-1}, T, V_{J+1}^N, J) \\
&= H(Z|J) - H(Z|U) - H(V) + H(V|U) \\
&\leq H(Z) - H(Z|U) - H(V) + H(V|U) \\
&= I(U; Z) - I(U; V), \tag{44}
\end{aligned}$$

where (a) follows from the fact that  $V_J$  is independent of  $J$ .

*Proof of  $p(V_J = v, J = i) = p(V_J = v)p(J = i)$ :* Since  $V^N$  is the output of a discrete memoryless source  $p_V(v)$ , we have

$$p(V_i = v) = P(V = v). \quad (45)$$

From <Part vi>, we know that the random variable  $J$  is independent of  $V^N$ , and therefore,

$$p(V_J = v, J = i) = p(V_i = v, J = i) = p(V_i = v)p(J = i) \stackrel{(1)}{=} P(V = v)p(J = i), \quad (46)$$

where (1) follows from (45).

On the other hand, the probability  $p(V_J = v)$  can be calculated as follows,

$$\begin{aligned} p(V_J = v) &= \sum_{i=1}^N p(V_J = v, J = i) = \sum_{i=1}^N p(V_i = v, J = i) \\ &\stackrel{(a)}{=} \sum_{i=1}^N P(V_i = v)p(J = i) \stackrel{(b)}{=} \sum_{i=1}^N P(V = v)p(J = i) \\ &= P(V = v) \sum_{i=1}^N p(J = i) = P(V = v), \end{aligned} \quad (47)$$

where (a) is from that  $J$  is independent of  $V^N$ , the formula (b) is from (45).

By using (46) and (47), it is easy to verify that  $V_J$  is independent of  $J$ , completing the proof. ■

Analogously, (36) is rewritten as follows,

$$\begin{aligned} \frac{1}{N} I(Y^N; S|T) &\leq \frac{1}{N} \sum_{i=1}^N (I(S; Y_i | Y^{i-1}, T, V_{i+1}^N, Z^{i-1}) - I(S; V_i | Y^{i-1}, V_{i+1}^N, T, Z^{i-1})) \\ &= \frac{1}{N} \sum_{i=1}^N (I(S; Y_i | Y^{i-1}, T, V_{i+1}^N, Z^{i-1}, J = i) - I(S; V_i | Y^{i-1}, V_{i+1}^N, T, Z^{i-1}, J = i)) \\ &= I(S; Y_J | Y^{J-1}, T, V_{J+1}^N, Z^{J-1}, J) - I(S; V_J | Y^{J-1}, T, V_{J+1}^N, Z^{J-1}, J) \\ &\stackrel{(a)}{=} I(K; Y|U, A) - I(K; V|U, A), \end{aligned} \quad (48)$$

where (a) follows from (40), (41), (42) and (43).

Similarly, (38) is rewritten as follows,

$$\frac{1}{N} I(S, T; Y^N) \leq I(U, K, A; Y) - I(U, K, A; V), \quad (49)$$

and (39) can be rewritten as follows,

$$\frac{1}{N} I(S; Z^N | T) = I(K; Z|U) - I(K; V|U). \quad (50)$$

Substituting (44), (48), (49), (50) into (30), (31), (32) and (33), Lemma 1 is proved. ■

The proof of Theorem 2 is completed.

## APPENDIX C

## SIZE CONSTRAINT OF THE RANDOM VARIABLES IN THEOREM 1

By using the support lemma (see [5], p.310), it suffices to show that the random variables  $U$  and  $K$  can be replaced by new ones, preserving the Markovity  $(U, K) \rightarrow (X, V) \rightarrow Y \rightarrow Z$  and the mutual information  $I(U; Z)$ ,  $I(U; V)$ ,  $I(K; Y|U)$ ,  $I(K; V|U)$ ,  $I(K; Z|U)$ , and furthermore, the ranges of the new  $U$  and  $K$  satisfy:  $\|\mathcal{U}\| \leq \|\mathcal{X}\| \|\mathcal{V}\| + 2$ ,  $\|\mathcal{K}\| \leq (\|\mathcal{X}\| \|\mathcal{V}\| + 2)^2$ . The proof is in the reminder of this section.

Let

$$\bar{p} = p_{XV}(x, v). \quad (51)$$

Define the following continuous scalar functions of  $\bar{p}$  :

$$f_{XV}(\bar{p}) = p_{XV}(x, v), f_Y(\bar{p}) = H(Y), f_Z(\bar{p}) = H(Z), f_V(\bar{p}) = H(V).$$

Since there are  $\|\mathcal{X}\| \|\mathcal{V}\| - 1$  functions of  $f_{XV}(\bar{p})$ , the total number of the continuous scalar functions of  $\bar{p}$  is  $\|\mathcal{X}\| \|\mathcal{V}\| + 2$ .

Let  $\bar{p}_{XV|U} = Pr\{X = x, V = v|U = u\}$ . With these distributions  $\bar{p}_{XV|U} = Pr\{X = x, V = v|U = u\}$ , we have

$$p_{XV}(x, v) = \sum_{u \in \mathcal{U}} p(U = u) f_{XV}(\bar{p}_{XV|U}), \quad (52)$$

$$I(U; Z) = f_Z(\bar{p}) - \sum_{u \in \mathcal{U}} p(U = u) f_Z(\bar{p}_{XV|U}), \quad (53)$$

$$I(U; V) = f_V(\bar{p}) - \sum_{u \in \mathcal{U}} p(U = u) f_V(\bar{p}_{XV|U}), \quad (54)$$

$$H(Y|U) = \sum_{u \in \mathcal{U}} p(U = u) f_Y(\bar{p}_{XV|U}), \quad (55)$$

According to the support lemma ([5], p.310), the random variable  $U$  can be replaced by new ones such that the new  $U$  takes at most  $\|\mathcal{X}\| \|\mathcal{V}\| + 2$  different values and the expressions (52), (53), (54) and (55) are preserved.

Once the alphabet of  $U$  is fixed, we apply similar arguments to bound the alphabet of  $K$ , see the following. Define  $\|\mathcal{X}\| \|\mathcal{V}\| + 2$  continuous scalar functions of  $\bar{p}_{XV}$  :

$$f_{XV}(\bar{p}_{XV}) = p_{XV}(x, v), f_Y(\bar{p}_{XV}) = H(Y), f_Z(\bar{p}_{XV}) = H(Z), f_V(\bar{p}_{XV}) = H(V),$$

where of the functions  $f_{XV}(\bar{p}_{XV})$ , only  $\|\mathcal{X}\| \|\mathcal{V}\| - 1$  are to be considered.

For every fixed  $u$ , let  $\bar{p}_{XV|K} = Pr\{X = x, V = v|K = k\}$ . With these distributions  $\bar{p}_{XV|K}$ , we have

$$Pr\{X = x, V = v|U = u\} = \sum_{k \in \mathcal{K}} Pr\{K = k|U = u\} f_{XV}(\bar{p}_{XV|K}), \quad (56)$$

$$I(K; Z|U) = H(Z|U = u) - \sum_{k \in \mathcal{K}} f_Z(\bar{p}_{XV|K}) Pr\{K = k|U = u\}, \quad (57)$$

$$I(K; V|U) = H(V|U = u) - \sum_{k \in \mathcal{K}} f_V(\bar{p}_{XV|K}) Pr\{K = k|U = u\}. \quad (58)$$

$$I(K; Y|U) = H(Y|U = u) - \sum_{k \in \mathcal{K}} f_Y(\bar{p}_{XV|K}) \Pr\{K = k|U = u\}. \quad (59)$$

By the support lemma ([5], p.310), for every fixed  $u$ , the size of the alphabet of the random variable  $K$  can not be larger than  $\|\mathcal{X}\|\|\mathcal{V}\| + 2$ , and therefore,  $\|\mathcal{K}\| \leq (\|\mathcal{X}\|\|\mathcal{V}\| + 2)^2$  is proved.

#### APPENDIX D

##### PROOF OF THEOREM 3

In this section, we will show that any triple  $(R_0, R_1, R_e) \in \mathcal{R}^{ci}$  is achievable. Superposition coding and Wyner's random binning techniques are used in the construction of the code-books.

Now the remainder of this section is organized as follows. The code construction is in Subsection D-A. The proof of achievability is given in Subsection D-B.

##### A. Code Construction

Since  $R_e \leq I(K; Y|U) - I(K; Z|U)$  and  $R_e \leq R_1 \leq I(K; Y|U)$ , it is sufficient to show that the triple  $(R_0, R_1, R_e = I(K; Y|U) - I(K; Z|U))$  is achievable, and note that this implies that  $R_1 \geq R_e = I(K; Y|U) - I(K; Z|U)$ .

Given a triple  $(R_0, R_1, R_e)$ , choose a joint probability mass function  $p_{U,K,V,X,Y,Z}(u, k, v, x, y, z)$  such that

$$0 \leq R_e \leq R_1,$$

$$R_0 \leq I(U; Z),$$

$$R_1 \leq I(K; Y|U),$$

$$R_e = I(K; Y|U) - I(K; Z|U).$$

The confidential message set  $\mathcal{S}$  and the common message set  $\mathcal{T}$  satisfy the following conditions:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \|\mathcal{S}\| = R_1 = I(K; Y|U) - \gamma_1, \quad (60)$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \|\mathcal{T}\| = R_0 = I(U; Z) - \gamma, \quad (61)$$

where  $\gamma$  and  $\gamma_1$  are fixed numbers and  $\gamma \geq 0$ ,

$$0 \leq \gamma_1 \leq^{(a)} I(K; Z|U). \quad (62)$$

Note that (a) is from  $R_1 \geq R_e = I(K; Y|U) - I(K; Z|U)$  and (60). Let  $\mathcal{S} = \{1, 2, \dots, 2^{NR_1}\}$  and  $\mathcal{T} = \{1, 2, \dots, 2^{NR_0}\}$ .

Code-book generation:

- **(Construction of  $U^N$ )**

For a given common message  $t$  ( $t \in \mathcal{T}$ ), generate a corresponding sequence  $u^N(t)$  i.i.d. according to the probability mass function  $p_U(u)$ .



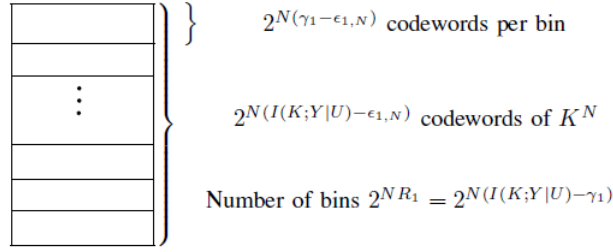


Fig. 5: Code-book construction for  $K^N$  in Theorem 3

- **(Construction of  $K^N$ )** Classical superposition coding and Wyner's random binning technique [15] are used in the construction of  $K^N$ , see Figure 5.

For the transmitted sequence  $u^N(t)$ , generate  $2^{N(I(K;Y|U) - \epsilon_{1,N})}$  ( $\epsilon_{1,N} \rightarrow 0$  as  $N \rightarrow \infty$ ) i.i.d. sequences  $k^N$ , according to the probability mass function  $p_{K|U}(k_i|u_i(t))$ . Distribute these sequences at random into  $2^{NR_1} = 2^{N(I(K;Y|U) - \gamma_1)}$  bins such that each bin contains  $2^{N(\gamma_1 - \epsilon_{1,N})}$  sequences. Index each bin by  $i \in \{1, 2, \dots, 2^{NR_1}\}$ .

Here note that the number of the sequences in every bin is upper bounded as follows.

$$\gamma_1 - \epsilon_{1,N} \stackrel{(a)}{\leq} I(K; Z|U) - \epsilon_{1,N}, \quad (63)$$

where (a) is from (62). This implies that

$$\lim_{N \rightarrow \infty} H(K^N|U^N, S, Z^N) = 0. \quad (64)$$

For a given confidential message  $s$  ( $s \in \mathcal{S}$ ), randomly choose a sequence  $k^N(u^N(t))$  in bin  $s$  for transmission.

- **(Construction of  $X^N$ )** The  $x^N$  is generated according to a new discrete memoryless channel (DMC) with inputs  $k^N$ ,  $u^N$ ,  $v^N$ , and output  $x^N$ . The transition probability of this new DMC is  $p_{X|U,K,V}(x|u,k,v)$ . The probability  $p_{X^N|U^N,K^N,V^N}(x^N|u^N,k^N,v^N)$  is calculated as follows.

$$p_{X^N|U^N,K^N,V^N}(x^N|u^N,k^N,v^N) = \prod_{i=1}^N p_{X|U,K,V}(x_i|u_i,k_i,v_i). \quad (65)$$

Decoding:

Receiver 2: Given a vector  $z^N \in \mathcal{Z}^N$ , try to find a sequence  $u^N(\hat{t})$  such that  $(u^N(\hat{t}), z^N) \in T_{UZ}^N(\epsilon)$ . If there exists a unique sequence, put out the corresponding  $\hat{t}$ . Otherwise, declare a decoding error.

Receiver 1: Given a vector  $y^N \in \mathcal{Y}^N$ , try to find a sequence  $u^N$  such that  $(u^N, y^N) \in T_{UY}^N(\epsilon)$ . If such a sequence does not exist, or there are more than one such sequence, declare a decoding error. Denote the corresponding sequence by  $u^N(\hat{t})$ , put out the corresponding index  $\hat{t}$ .

After decoding  $u^N(\hat{t})$  and  $\hat{t}$ , try to find a sequence  $k^N(u^N(\hat{t}))$  such that  $(k^N(u^N(\hat{t})), y^N) \in T_{KY|U}^N(\epsilon)$ . If there exist sequences with the same index of the bin  $\hat{s}$ , put out the corresponding  $\hat{s}$ . Otherwise, declare a decoding error.

### B. Proof of Achievability

By using the above definitions, it is easy to verify that  $\lim_{N \rightarrow \infty} \frac{\log \|\mathcal{T}\|}{N} = R_0$  and  $\lim_{N \rightarrow \infty} \frac{\log \|\mathcal{S}\|}{N} = R_1$ .

Then, note that the above encoding and decoding scheme is similar to the one used in [14]. Hence, by similar arguments as in [14], it is easy to show that  $P_{e1} \leq \epsilon$  and  $P_{e2} \leq \epsilon$ , and the proof is omitted here. It remains to show that  $\lim_{N \rightarrow \infty} \Delta \geq R_e$ , see the following.

$$\begin{aligned}
\lim_{N \rightarrow \infty} \Delta &= \lim_{N \rightarrow \infty} \frac{1}{N} H(S|Z^N) \\
&\geq \lim_{N \rightarrow \infty} \frac{1}{N} H(S|Z^N, U^N) \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} (H(S, Z^N, U^N) - H(Z^N, U^N)) \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} (H(S, Z^N, U^N, K^N) - H(K^N|Z^N, U^N, S) - H(Z^N, U^N)) \\
&\stackrel{(a)}{=} \lim_{N \rightarrow \infty} \frac{1}{N} (H(Z^N|U^N, K^N) + H(U^N, K^N, S) - H(K^N|Z^N, U^N, S) - H(Z^N, U^N)) \\
&\stackrel{(b)}{=} \lim_{N \rightarrow \infty} \frac{1}{N} (H(Z^N|U^N, K^N) + H(U^N, K^N) - H(K^N|Z^N, U^N, S) - H(Z^N, U^N)) \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} (H(K^N|U^N) - H(K^N|Z^N, U^N, S) - I(Z^N; K^N|U^N)) \\
&\geq \lim_{N \rightarrow \infty} \frac{1}{N} (H(K^N|U^N) - H(K^N|U^N, Y^N) - H(K^N|Z^N, U^N, S) - I(Z^N; K^N|U^N)) \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} (I(Y^N; K^N|U^N) - H(K^N|Z^N, U^N, S) - I(Z^N; K^N|U^N)) \\
&\stackrel{(c)}{=} \lim_{N \rightarrow \infty} \frac{1}{N} (NI(Y; K|U) - H(K^N|Z^N, U^N, S) - NI(K; Z|U)) \\
&\stackrel{(d)}{=} \lim_{N \rightarrow \infty} \frac{1}{N} (NI(Y; K|U) - NI(K; Z|U)) \\
&= I(Y; K|U) - I(K; Z|U) = R_e,
\end{aligned} \tag{66}$$

where (a) is from  $S \rightarrow (U^N, K^N) \rightarrow Z^N$ , (b) is from  $H(S|U^N, K^N) = 0$ , (c) is from that  $V^N, U^N, K^N$  and  $X^N$  are i.i.d. generated random vectors, and the channels are discrete memoryless, and (d) is from (64).

Thus,  $\lim_{N \rightarrow \infty} \Delta \geq R_e$  is proved, and the proof of Theorem 3 is completed.

## APPENDIX E

### PROOF OF THEOREM 4

In this section, we prove Theorem 4: all the achievable  $(R_0, R_1, R_e)$  triples are contained in the set  $\mathcal{R}^{(co)}$ . Suppose  $(R_0, R_1, R_e)$  is achievable, i.e., for any given  $\epsilon > 0$ , there exists a channel encoder-decoder  $(N, \Delta, P_{e1}, P_{e2})$  such that

$$\lim_{N \rightarrow \infty} \frac{\log \|\mathcal{T}\|}{N} = R_0, \lim_{N \rightarrow \infty} \frac{\log \|\mathcal{S}\|}{N} = R_1, \lim_{N \rightarrow \infty} \Delta \geq R_e, P_{e1} \leq \epsilon, P_{e2} \leq \epsilon.$$

Then we will show the existence of random variables  $(U, K, A) \rightarrow (X, V) \rightarrow Y \rightarrow Z$  such that

$$0 \leq R_e \leq R_1, \tag{67}$$

$$R_0 \leq I(U; Z), \quad (68)$$

$$R_1 \leq I(K; Y|U), \quad (69)$$

$$R_e \leq I(K; Y|U) - I(K; Z|A). \quad (70)$$

The formula (67) is from

$$R_e \leq \lim_{N \rightarrow \infty} \Delta = \lim_{N \rightarrow \infty} \frac{1}{N} H(S|Z^N) \leq \lim_{N \rightarrow \infty} \frac{1}{N} H(S) = R_1.$$

Since the model of Figure 2 with causal side information is a special case of the model of Figure 2 with noncausal side information, the formulas (68), (69) and (70) are obtained from (71), (72) and (73), respectively, see the following.

*Proof of (68):* The parameter  $R_0$  of (68) can be written as follows,

$$\begin{aligned} R_0 &= \lim_{N \rightarrow \infty} \frac{H(T)}{N} \\ &\stackrel{(a)}{\leq} \lim_{N \rightarrow \infty} \left( \frac{1}{N} \delta(P_{e2}) + \frac{1}{N} \sum_{i=1}^N (H(Z_i) - H(Z_i|Z^{i-1}, T, V_{i+1}^N)) \right) \\ &\leq \lim_{N \rightarrow \infty} \left( \frac{1}{N} \delta(P_{e2}) + \frac{1}{N} \sum_{i=1}^N (H(Z_i) - H(Z_i|Y^{i-1}, Z^{i-1}, T, V_{i+1}^N)) \right) \\ &\stackrel{(b)}{\leq} \lim_{N \rightarrow \infty} \left( \frac{1}{N} \delta(\epsilon) + H(Z) - H(Z|U) \right) \\ &\stackrel{(c)}{=} I(U; Z), \end{aligned} \quad (71)$$

where (a) follows from (30) and (34), and  $V_i$  is independent of  $(Z^{i-1}, T, V_{i+1}^N)$ , the formula (b) is from the definitions  $U \triangleq (T, V_{J+1}^N, Y^{J-1}, Z^{J-1}, J)$ ,  $Z \triangleq Z_J$ , and the formula (c) follows from  $\epsilon \rightarrow 0$ . Thus, the proof of (68) is completed. ■

*Proof of (69):* The parameter  $R_1$  of (69) can be written follows,

$$\begin{aligned} R_1 &= \lim_{N \rightarrow \infty} \frac{H(S)}{N} \\ &\stackrel{(1)}{\leq} \lim_{N \rightarrow \infty} \left( \frac{1}{N} (I(S; Y^N|T) + \frac{1}{N} \delta(P_{e1})) \right) \\ &\leq \lim_{N \rightarrow \infty} \left( \frac{1}{N} I(S; Y^N|T) + \frac{1}{N} \delta(\epsilon) \right) \\ &\stackrel{(2)}{\leq} \lim_{N \rightarrow \infty} \left( \frac{1}{N} \sum_{i=1}^N I(S; Y_i|T, V_{i+1}^N, Y^{i-1}, Z^{i-1}) + \frac{1}{N} \delta(\epsilon) \right) \\ &\stackrel{(3)}{\leq} I(K; Y|U), \end{aligned} \quad (72)$$

where (1) follows from (31), the formula (2) is from (36) and the fact that  $I(S; V_i|T, V_{i+1}^N, Y^{i-1}, Z^{i-1}) = 0$ , the formula (3) is from the definitions  $U \triangleq (T, V_{J+1}^N, Y^{J-1}, Z^{J-1}, J)$ ,  $K \triangleq S$ ,  $Y \triangleq Y_J$ , ( $J$  is independent of  $S$ ,  $T$ ,  $X^N$ ,  $V^N$ ,  $Y^N$  and  $Z^N$ , and  $J$  is uniformly distributed over  $\{1, 2, \dots, N\}$ ) and the fact that  $\epsilon \rightarrow 0$ . Thus, the formula (69) is proved. ■

*Proof of (70):* The parameter  $R_e$  of (70) satisfies

$$\begin{aligned}
R_e &= \lim_{N \rightarrow \infty} \frac{H(S|Z^N)}{N} \\
&\stackrel{(a)}{\leq} \lim_{N \rightarrow \infty} \left( \frac{1}{N} I(S; Y^N|T) - \frac{1}{N} I(S; Z^N|T) + \frac{1}{N} \delta(P_{e1}) + \frac{1}{N} \delta(P_{e2}) \right) \\
&\stackrel{(b)}{\leq} \lim_{N \rightarrow \infty} \left( \frac{1}{N} \sum_{i=1}^N (I(S; Y_i|T, V_{i+1}^N, Y^{i-1}, Z^{i-1}) - \right. \\
&\quad \left. I(S; Z_i|Z^{i-1}, T, V_{i+1}^N)) + \frac{1}{N} \delta(P_{e1}) + \frac{1}{N} \delta(P_{e2}) \right) \\
&\stackrel{(c)}{=} I(K; Y|U) - I(K; Z|A), \tag{73}
\end{aligned}$$

where (a) follows from (33), the formula (b) is from (36) and (39), and the fact that  $V_i$  is independent of  $(Z^{i-1}, Y^{i-1}, T, S, V_{i+1}^N)$ , the formula (c) is from the definitions  $U \triangleq (T, V_{J+1}^N, Y^{J-1}, Z^{J-1}, J)$ ,  $K \triangleq S$ ,  $A \triangleq (T, V_{J+1}^N, Z^{J-1}, J)$ ,  $Y \triangleq Y_J$ ,  $Z \triangleq Z_J$ , and the fact that  $P_{e1}, P_{e2} \leq \epsilon \rightarrow 0$ . Thus, the proof of (70) is completed.  $\blacksquare$

The proof of Theorem 4 is completed.

## APPENDIX F

### PROOF OF THEOREM 5

In this section, we will show that any triple  $(R_0, R_1, R_e) \in \mathcal{R}^{nfi}$  is achievable. Superposition coding, Gel'fand-Pinsker's binning, block Markov coding and Ahlswede-Cai's secret key on feedback [1] are used in the construction of the code-books. In addition, the encoding and decoding scheme for Theorem 5 can be also viewed as a combination of Steinberg's method [14], block Markov coding and Ahlswede-Cai's secret key on feedback [1].

Now the remainder of this section is organized as follows. The code construction is in Subsection F-A. The proof of achievability is given in Subsection F-B.

#### A. Code Construction

Given a triple  $(R_0, R_1, R_e)$ , choose a joint probability mass function  $p_{U,K,V,X,Y,Z}(u, k, v, x, y, z)$  such that

$$0 \leq R_e \leq R_1,$$

$$R_0 \leq I(U; Z) - I(U; V),$$

$$R_1 \leq I(K; Y|U) - I(K; V|U),$$

$$R_e \leq H(Y|Z).$$

The confidential message set  $\mathcal{S}$  and the common message set  $\mathcal{T}$  satisfy the following conditions:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \|\mathcal{S}\| = R_1 = I(K; Y|U) - I(K; V|U) - \gamma_1, \tag{74}$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \|\mathcal{T}\| = R_0 = I(U; Z) - I(U; V) - \gamma, \tag{75}$$

where  $\gamma$  and  $\gamma_1$  are fixed positive real numbers. Let  $\mathcal{S} = \{1, 2, \dots, 2^{NR_1}\}$  and  $\mathcal{T} = \{1, 2, \dots, 2^{NR_0}\}$ .

We use the block Markov coding method. The random vectors  $U^N$ ,  $K^N$ ,  $V^N$ ,  $X^N$ ,  $Y^N$  and  $Z^N$  consist of  $n$  blocks of length  $N$ . The common message for  $n$  blocks is  $T^n \triangleq (T_1, \dots, T_n)$ , which is composed of  $n$  i.i.d. random variables uniformly distributed over  $\mathcal{T}$ . The confidential message for  $n$  blocks is  $S^n \triangleq (S_2, S_3, \dots, S_n)$ , where  $S_i$  ( $2 \leq i \leq n$ ) are i.i.d. random variables uniformly distributed over  $\mathcal{S}$ . Note that in the first block, there is no  $S_1$ .

Let  $\tilde{Z}_i$  ( $1 \leq i \leq n$ ) be the output of channel 2 for block  $i$ ,  $Z^n = (\tilde{Z}_1, \dots, \tilde{Z}_n)$ ,  $Z^j = (\tilde{Z}_1, \dots, \tilde{Z}_{j-1}, \tilde{Z}_{j+1}, \dots, \tilde{Z}_n)$  ( $1 \leq j \leq n$ ). Similarly,  $Y^n = (\tilde{Y}_1, \dots, \tilde{Y}_n)$ , and  $\tilde{Y}_i$  ( $1 \leq i \leq n$ ) is the output of channel 1 for block  $i$ . The specific values of the above random vectors are denoted by lower case letters.

Code-book generation:

- **(Construction of  $U^N$ )**

The construction of  $U^N$  for each block is exactly the same as that in the proof of Theorem 1, see the following. For the  $i$ -th block ( $1 \leq i \leq n$ ), generate  $2^{N(I(U;Z)-\epsilon_{1,N})}$  ( $\epsilon_{1,N} \rightarrow 0$  as  $N \rightarrow \infty$ ) i.i.d. sequences  $u^N$ , according to the probability mass function  $p_U(u)$ . Distribute these sequences at random into  $2^{NR_0} = 2^{N(I(U;Z)-I(U;V)-\gamma)}$  bins such that each bin contains  $2^{N(I(U;V)+\gamma-\epsilon_{1,N})}$  sequences. Index each bin by  $t_i \in \{1, 2, \dots, 2^{NR_0}\}$ . For a given common message  $t_i$  ( $t_i \in \mathcal{T}$ ) and side information  $v^N$ , try to find a sequence in bin  $t_i$   $\{u^N(t_i, 1), u^N(t_i, 2), \dots, u^N(t_i, 2^{N(I(U;V)+\gamma-\epsilon_{1,N})})\}$  that is jointly typical with  $v^N$ , say  $u^N(t_i, j^*)$ , i.e.,  $(u^N(t_i, j^*), v^N) \in T_{UV}^N(\epsilon_1)$ . If multiple such sequences in bin  $t_i$  exist, choose the one with the smallest  $j^*$ . If no such  $j^*$  exists, then declare an encoding error.

- **(Construction of  $K^N$ )**

Superposition coding, Gel'fand-Pinsker's binning, block Markov coding and Ahlswede-Cai's secret key on feedback [1] are used in the construction of  $K^N$ , see Figure 6.

For the transmitted sequence  $u^N(t_i, j^*)$ , generate  $2^{N(I(K;Y|U)-\epsilon_{2,N})}$  ( $\epsilon_{2,N} \rightarrow 0$  as  $N \rightarrow \infty$ ) i.i.d. sequences  $k^N$ , according to the probability mass function  $p_{K|U}(k_i|u_i)$ . Distribute these sequences at random into  $2^{NR_1} = 2^{N(I(K;Y|U)-I(K;V|U)-\gamma_1)}$  bins such that each bin contains  $2^{N(I(K;V|U)+\gamma_1-\epsilon_{2,N})}$  sequences. Index each bin by  $l \in \{1, 2, \dots, 2^{NR_1}\}$ .

In the first block, for a given side information  $v^N$ , try to find a  $k^N(u^N(t_1, j^*))$  generated by  $u^N(t_1, j^*)$  such that  $(k^N(u^N(t_1, j^*)), v^N) \in T_{KV|U}^N(\epsilon)$ . If multiple such sequences exist, randomly choose one for transmission. If no such sequence exists, declare an encoding error.

For the  $i$ -th block ( $2 \leq i \leq n$ ), firstly we generate a mapping  $g_f : \mathcal{Y}^N \rightarrow \mathcal{S}$  (note that  $\|\mathcal{Y}\|^N \geq \|\mathcal{S}\|$ ). Define a random variable  $K_i^* = g_f(\tilde{Y}_{i-1})$  ( $2 \leq i \leq n$ ), which is uniformly distributed over  $\mathcal{S}$ , and  $K_i^*$  is independent of  $S_i$ . Reveal the mapping  $g_f$  to both receivers and the transmitter.

Then, when the transmitter receives the output  $\tilde{y}_{i-1}$  of the  $i-1$ -th block, he computes  $k_i^* = g_f(\tilde{y}_{i-1}) \in \mathcal{S}$ . For a given  $s_i$ , the transmitter chooses a sequence  $k^N(u^N(t_i, j^*))$  from the bin  $s_i \oplus k_i^*$  (where  $\oplus$  is the modulo addition over  $\mathcal{S}$ ) such that  $(k^N(u^N(t_i, j^*)), v^N) \in T_{KV|U}^N(\epsilon)$ . If multiple such sequences in bin  $s_i \oplus k_i^*$  exist, choose the one with the smallest index in the bin. If no such sequence exists, declare an encoding error.

- **(Construction of  $X^N$ )**

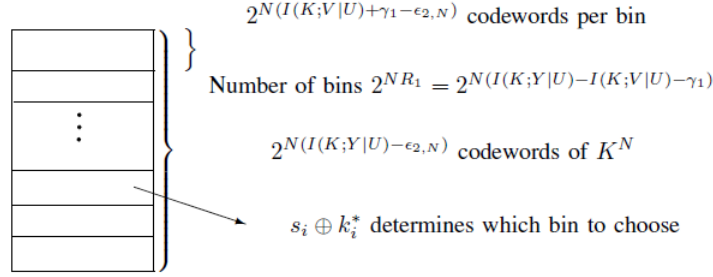


Fig. 6: Code-book construction for  $K^N$  in Theorem 5, where  $s_i$  and  $k_i^*$  are the confidential message and the secret key for block  $i$ , respectively.

For each block, the  $x^N$  is generated according to a new discrete memoryless channel (DMC) with inputs  $k^N$ ,  $u^N$ ,  $v^N$ , and output  $x^N$ . The transition probability of this new DMC is  $p_{X|U,K,V}(x|u, k, v)$ . The probability  $p_{X^N|U^N, K^N, V^N}(x^N|u^N, k^N, v^N)$  is calculated as follows.

$$p_{X^N|U^N, K^N, V^N}(x^N|u^N, k^N, v^N) = \prod_{i=1}^N p_{X|U,K,V}(x_i|u_i, k_i, v_i). \quad (76)$$

Decoding:

Receiver 2: For each block, given a vector  $z^N \in \mathcal{Z}^N$ , try to find a sequence  $u^N(\hat{t}, \hat{i})$  such that  $(u^N(\hat{t}, \hat{i}), z^N) \in T_{UZ}^N(\epsilon_3)$ . If there exist sequences with the same  $\hat{t}$ , put out the corresponding  $\hat{t}$ . Otherwise, i.e., if no such sequence exists or multiple sequences have different message indices, declare a decoding error.

Receiver 1: For the first block, given a vector  $\tilde{y}_1 \in \mathcal{Y}^N$ , try to find a sequence  $u^N$  such that  $(u^N, \tilde{y}_1) \in T_{UY}^N(\epsilon_4)$ . If such a sequence does not exist, or there are more than one such sequence, declare a decoding error. Denote the corresponding sequence by  $u^N(\hat{t}_1, \hat{j})$ , and put out the corresponding index  $\hat{t}_1$ .

For block  $i$  ( $2 \leq i \leq n$ ), given a vector  $\tilde{y}_i \in \mathcal{Y}^N$ , try to find a sequence  $u^N$  such that  $(u^N, \tilde{y}_i) \in T_{UY}^N(\epsilon_4)$ . If such a sequence does not exist, or there are more than one such sequence, declare a decoding error. Denote the corresponding sequence by  $u^N(\hat{t}_i, \hat{j})$ , put out the corresponding index  $\hat{t}_i$ .

After decoding  $u^N(\hat{t}_i, \hat{j})$  and  $\hat{t}_i$ , try to find a sequence  $k^N(u^N(\hat{t}_i, \hat{j}))$  such that  $(k^N(u^N(\hat{t}_i, \hat{j})), \tilde{y}_i) \in T_{KY|U}^N(\epsilon_5)$ . If there exist sequences with the same index of the bin  $\hat{s}_i \oplus k_i^*$ , put out the corresponding  $\hat{s}_i \oplus k_i^*$ . Otherwise, declare a decoding error. Finally, since receiver 1 knows  $k_i^*$  ( $k_i^* = g_f(\tilde{y}_{i-1})$ ), put out the corresponding  $\hat{s}_i$  from  $\hat{s}_i \oplus k_i^*$ .

### B. Proof of Achievability

The rate of the confidential message  $S^n$  is defined as

$$R_1^* \triangleq \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{H(S^n)}{nN}, \quad (77)$$

and it satisfies

$$\begin{aligned}
R_1^* &= \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{H(S^n)}{nN} \\
&= \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{\sum_{i=2}^n H(S_i)}{nN} \\
&= \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{(n-1)H(S)}{nN} \\
&= R_1.
\end{aligned} \tag{78}$$

Similarly, the rate of the common messages  $T^n$  satisfies

$$\begin{aligned}
R_0^* &\triangleq \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{H(T^n)}{nN} \\
&= R_0.
\end{aligned} \tag{79}$$

In addition, note that the encoding and decoding scheme for Theorem 5 is exactly the same as that in [14], except that the transmitted message for receiver 1 is  $s \oplus k^*$ . Since receiver 1 knows  $k^*$ , the decoding scheme for Theorem 5 is in fact the same as that in [14]. Hence, we omit the proof of  $P_{e1} \leq \epsilon$  and  $P_{e2} \leq \epsilon$  here.

It remains to show that  $\lim_{N \rightarrow \infty} \Delta \geq R_e$ , see the following.

Since the confidential message  $S$  is encrypted by  $S \oplus K^*$ , the equivocation about  $S$  is equivalent to the equivocation about  $K^*$ . There are two ways for receiver 2 to obtain the secret key  $k^*$ . One way is that he tries to guess the  $k^*$  from its alphabet  $\mathcal{S}$ . The other way is that he tries to guess the feedback  $y^N$  ( $y^N$  is the output of channel 1 for the previous block, and  $k^* = g_f(y^N)$ ) from the conditional typical set  $T_{[Y|Z]}^N(\delta)$ , and this is because for a given  $z^N$  and sufficiently large  $N$ ,  $\Pr\{(y^N \notin T_{[Y|Z]}^N(\delta))\} \rightarrow 0$ . Note that there are  $2^{NH(Y|Z)}$  sequences  $y^N \in T_{[Y|Z]}^N(\delta)$  when  $N \rightarrow \infty$  and  $\delta \rightarrow 0$ . Therefore, the equivocation about  $S$  is  $\min\{\frac{\log \|\mathcal{S}\|}{N} = R_1, H(Y|Z)\}$ , and note that  $R_1 \geq R_e$  and  $H(Y|Z) \geq R_e$ , then  $\lim_{N \rightarrow \infty} \Delta \geq R_e$  is obtained.

The details about the proof are as follows.

First, we will show that  $K_i^* \oplus S_i$  is independent of  $K_i^*$  and  $S_i$ , and this is used in the proof of  $\lim_{N \rightarrow \infty} \Delta \geq R_e$ .

Since  $K_i^*$  is independent of  $S_i$  ( $2 \leq i \leq n$ ), and all of them are uniformly distributed over  $\mathcal{S}$ , the fact that  $K^* \oplus S_i$  is independent of  $K^*$  and  $S_i$  is proved by the following (80) and (81).

$$\begin{aligned}
\Pr\{K_i^* \oplus S_i = a\} &= \sum_{k_i^* \in \mathcal{S}} \Pr\{K_i^* \oplus S_i = a, K_i^* = k_i^*\} \\
&= \sum_{k_i^* \in \mathcal{S}} \Pr\{S_i = a \ominus k_i^*, K_i^* = k_i^*\} \\
&= \sum_{k_i^* \in \mathcal{S}} \Pr\{S_i = a \ominus k_i^*\} \Pr\{K_i^* = k_i^*\} \\
&= \frac{1}{\|\mathcal{S}\|}.
\end{aligned} \tag{80}$$

$$\begin{aligned}
Pr\{K_i^* \oplus S_i = a, K_i^* = k_i^*\} &= Pr\{S_i = a \ominus k_i^*, K_i^* = k_i^*\} \\
&= Pr\{S_i = a \ominus k_i^*\} Pr\{K_i^* = k_i^*\} \\
&= \frac{1}{\|\mathcal{S}\|^2}.
\end{aligned} \tag{81}$$

Then,  $\lim_{N \rightarrow \infty} \Delta \geq R_e$  is proved by the following (82).

$$\begin{aligned}
\lim_{N \rightarrow \infty} \Delta &\triangleq \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{H(S^n | Z^n)}{nN} \\
&= \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{\sum_{i=2}^n H(S_i | S^{i-1}, Z^n)}{nN} \\
&\stackrel{(a)}{=} \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{\sum_{i=2}^n H(S_i | \tilde{Z}_i, \tilde{Z}_{i-1})}{nN} \\
&\geq \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{\sum_{i=2}^n H(S_i | \tilde{Z}_i, \tilde{Z}_{i-1}, S_i \oplus K_i^*)}{nN} \\
&\stackrel{(b)}{=} \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{\sum_{i=2}^n H(S_i | \tilde{Z}_{i-1}, S_i \oplus K_i^*)}{nN} \\
&= \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{\sum_{i=2}^n H(K_i^* | \tilde{Z}_{i-1}, S_i \oplus K_i^*)}{nN} \\
&\stackrel{(c)}{=} \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{\sum_{i=2}^n H(K_i^* | \tilde{Z}_{i-1})}{nN} \\
&\stackrel{(d)}{=} \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{\sum_{i=2}^n \min\{NH(Y|Z), NR_1\}}{nN} \\
&= \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{(n-1) \min\{NR_1, NH(Y|Z)\}}{nN} \\
&= \min\{R_1, H(Y|Z)\} \\
&\geq R_e,
\end{aligned} \tag{82}$$

where (a) is from  $S_i \rightarrow (\tilde{Z}_i, \tilde{Z}_{i-1}) \rightarrow (S^{i-1}, \tilde{Z}^{i-2}, \tilde{Z}_{i+1}^n)$  is a Markov chain, (b) is from  $S_i \rightarrow (S_i \oplus K_i^*, \tilde{Z}_{i-1}) \rightarrow \tilde{Z}_i$  is a Markov chain, (c) follows from the fact that  $K_i^* \oplus S_i$  is independent of  $K_i^*$  and  $\tilde{Z}_{i-1}$ , and (d) is from the fact that receiver 2 can guess the specific vector  $\tilde{Y}_{i-1}$  (corresponding to the key  $K_i^*$ ) from the conditional typical set  $T_{[Y|Z]}^N(\delta)$ , and  $K_i^*$  is uniformly distributed over  $\mathcal{S}$  ( $K_i^*$  is the key used in block  $i$ ).

Thus,  $\lim_{N \rightarrow \infty} \Delta \geq R_e$  is proved, and the proof of Theorem 5 is completed.

## APPENDIX G

### PROOF OF THEOREM 6

In this section, we prove Theorem 6: all the achievable  $(R_0, R_1, R_e)$  triples are contained in the set  $\mathcal{R}^{(nfo)}$ . Suppose  $(R_0, R_1, R_e)$  is achievable, i.e., for any given  $\epsilon > 0$ , there exists a channel encoder-decoder  $(N, \Delta, P_{e1}, P_{e2})$  such that

$$\lim_{N \rightarrow \infty} \frac{\log \|\mathcal{T}\|}{N} = R_0, \lim_{N \rightarrow \infty} \frac{\log \|\mathcal{S}\|}{N} = R_1, \lim_{N \rightarrow \infty} \Delta \geq R_e, P_{e1} \leq \epsilon, P_{e2} \leq \epsilon.$$



Then we will show the existence of random variables  $(U, K, A) \rightarrow (X, V) \rightarrow Y \rightarrow Z$  such that

$$0 \leq R_e \leq R_1, \quad (83)$$

$$R_0 \leq I(U; Z) - I(U; V), \quad (84)$$

$$R_1 \leq I(K; Y|U, A) - I(K; V|U, A), \quad (85)$$

$$R_0 + R_1 \leq I(U, K, A; Y) - I(U, K, A; V), \quad (86)$$

$$R_e \leq H(Y|Z). \quad (87)$$

The proof of (83), (84), (85) and (86) are the same as those in the proof of Theorem 2, and therefore, we omit it here. It remains to prove (87), see the following.

$$\begin{aligned}
\frac{1}{N}H(S|Z^N) &\stackrel{(a)}{\leq} \frac{1}{N}H(S|Z^N) + \frac{1}{N}\delta(P_{e1}) - \frac{1}{N}H(S|Y^N, Z^N) \\
&= \frac{1}{N}I(S; Y^N|Z^N) + \frac{1}{N}\delta(P_{e1}) \\
&\leq \frac{1}{N}H(Y^N|Z^N) + \frac{1}{N}\delta(P_{e1}) \\
&\leq \frac{1}{N} \sum_{i=1}^N H(Y_i|Z_i) + \frac{1}{N}\delta(P_{e1}) \\
&\stackrel{(b)}{=} \frac{1}{N} \sum_{i=1}^N H(Y_i|Z_i, J=i) + \frac{1}{N}\delta(P_{e1}) \\
&= H(Y_J|Z_J, J) + \frac{1}{N}\delta(P_{e1}) \\
&\stackrel{(c)}{=} H(Y|Z, J) + \frac{1}{N}\delta(P_{e1}) \\
&\leq H(Y|Z) + \frac{1}{N}\delta(\epsilon), \quad (88)
\end{aligned}$$

where (a) is from Fano's inequality, (b) is from  $J$  is independent of  $Y^N$  and  $Z^N$ , and (c) is from  $Y \triangleq Y_J$  and  $Z \triangleq Z_J$ .

By using  $\lim_{N \rightarrow \infty} \Delta \geq R_e$ ,  $\Delta = \frac{1}{N}H(S|Z^N)$ , (88) and letting  $\epsilon \rightarrow 0$ , the formula (87) is proved.

The proof of Theorem 6 is completed.

## APPENDIX H

### PROOF OF THEOREM 7

#### A. Converse Part of Theorem 7

The converse proof of Theorem 7 is exactly the same as the proof of Theorem 4, except that  $R_e \leq H(Y|Z)$ . Note that the proof of  $R_e \leq H(Y|Z)$  is the same as that in the proof of Theorem 6, and hence we omit the proof here.

### B. Direct Part of Theorem 7

In this section, we will show that any triple  $(R_0, R_1, R_e) \in \mathcal{R}^{cf}$  is achievable.

1) *Code Construction:* Given a triple  $(R_0, R_1, R_e)$ , choose a joint probability mass function  $p_{U,K,V,X,Y,Z}(u, k, v, x, y, z)$  such that

$$0 \leq R_e \leq R_1,$$

$$R_0 \leq I(U; Z),$$

$$R_1 \leq I(K; Y|U),$$

$$R_e \leq H(Y|Z).$$

The confidential message set  $\mathcal{S}$  and the common message set  $\mathcal{T}$  satisfy the following conditions:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \|\mathcal{S}\| = R_1 = I(K; Y|U) - \gamma_1, \quad (89)$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \|\mathcal{T}\| = R_0 = I(U; Z) - \gamma, \quad (90)$$

where  $\gamma$  and  $\gamma_1$  are fixed positive numbers. Let  $\mathcal{S} = \{1, 2, \dots, 2^{NR_1}\}$  and  $\mathcal{T} = \{1, 2, \dots, 2^{NR_0}\}$ .

We use the block Markov coding method. The random vectors  $U^N, K^N, V^N, X^N, Y^N$  and  $Z^N$  consist of  $n$  blocks of length  $N$ . The common message for  $n$  blocks is  $T^n \triangleq (T_1, \dots, T_n)$ , which is composed of  $n$  i.i.d. random variables uniformly distributed over  $\mathcal{T}$ . The confidential message for  $n$  blocks is  $S^n \triangleq (S_2, S_3, \dots, S_n)$ , where  $S_i$  ( $2 \leq i \leq n$ ) are i.i.d. random variables uniformly distributed over  $\mathcal{S}$ . Note that in the first block, there is no  $S_1$ .

Let  $\tilde{Z}_i$  ( $1 \leq i \leq n$ ) be the output of channel 2 for block  $i$ ,  $Z^n = (\tilde{Z}_1, \dots, \tilde{Z}_n)$ ,  $Z^j = (\tilde{Z}_1, \dots, \tilde{Z}_{j-1}, \tilde{Z}_{j+1}, \dots, \tilde{Z}_n)$  ( $1 \leq j \leq n$ ). Similarly,  $Y^n = (\tilde{Y}_1, \dots, \tilde{Y}_n)$ , and  $\tilde{Y}_i$  ( $1 \leq i \leq n$ ) is the output of channel 1 for block  $i$ . The specific values of the above random vectors are denoted by lower case letters.

Code-book generation:

- **(Construction of  $U^N$ )**

The construction of  $U^N$  for each block is exactly the same as that in the proof of Theorem 3, see the following. For the  $i$ -th block ( $1 \leq i \leq n$ ), given a common message  $t_i$  ( $t_i \in \mathcal{T}$ ), generate a corresponding  $u^N(t_i)$  i.i.d. according to the probability mass function  $p_U(u)$ .

- **(Construction of  $K^N$ )**

In the first block, for a given  $u^N(t_1)$ , generate a corresponding sequence  $k^N$  i.i.d. according to the probability mass function  $p_{K|U}(k_i|u_i)$ .

In the  $i$ -th block ( $2 \leq i \leq n$ ), firstly we generate a mapping  $g_f : \mathcal{Y}^N \rightarrow \mathcal{S}$ . Define a random variable  $K_i^* = g_f(\tilde{Y}_{i-1})$  ( $2 \leq i \leq n$ ), which is uniformly distributed over  $\mathcal{S}$ , and  $K_i^*$  is independent of  $S_i$ . Reveal the mapping  $g_f$  to both receivers and the transmitter. Then, when the transmitter receives the output  $\tilde{y}_{i-1}$  of the  $i$ -1-th block, he computes  $k_i^* = g_f(\tilde{y}_{i-1}) \in \mathcal{S}$ .

Given the transmitted sequence  $u^N(t_i)$  and the encrypted confidential message  $s_i \oplus k_i^*$  (where  $\oplus$  is the modulo addition over  $\mathcal{S}$ ), generate a corresponding sequence  $k^N$  i.i.d. according to the probability mass function  $p_{K|U}(k_i|u_i)$ . Index  $k^N$  by  $s_i \oplus k_i^* \in \mathcal{S}$ .

• **(Construction of  $X^N$ )**

For each block, the  $x^N$  is generated according to a new discrete memoryless channel (DMC) with inputs  $k^N$ ,  $u^N$ ,  $v^N$ , and output  $x^N$ . The transition probability of this new DMC is  $p_{X|U,K,V}(x|u, k, v)$ . The probability  $p_{X^N|U^N, K^N, V^N}(x^N|u^N, k^N, v^N)$  is calculated as follows.

$$p_{X^N|U^N, K^N, V^N}(x^N|u^N, k^N, v^N) = \prod_{i=1}^N p_{X|U,K,V}(x_i|u_i, k_i, v_i). \quad (91)$$

Decoding:

Receiver 2: For each block, given a vector  $z^N \in \mathcal{Z}^N$ , try to find a sequence  $u^N(\hat{t})$  such that  $(u^N(\hat{t}), z^N) \in T_{UZ}^N(\epsilon_3)$ . If there exists a unique sequence, put out the corresponding  $\hat{t}$ . Otherwise, declare a decoding error.

Receiver 1: For the first block, given a vector  $\tilde{y}_1 \in \mathcal{Y}^N$ , try to find a sequence  $u^N$  such that  $(u^N, \tilde{y}_1) \in T_{UY}^N(\epsilon_4)$ . If such a sequence does not exist, or there are more than one such sequence, declare a decoding error. Denote the corresponding sequence by  $u^N(\hat{t}_1)$ , put out the corresponding index  $\hat{t}_1$ .

For block  $i$  ( $2 \leq i \leq n$ ), given a vector  $\tilde{y}_i \in \mathcal{Y}^N$ , try to find a sequence  $u^N$  such that  $(u^N, \tilde{y}_i) \in T_{UY}^N(\epsilon_4)$ . If such a sequence does not exist, or there are more than one such sequence, declare a decoding error. Denote the corresponding sequence by  $u^N(\hat{t}_i)$ , put out the corresponding index  $\hat{t}_i$ .

After decoding  $u^N(\hat{t}_i)$  and  $\hat{t}_i$ , try to find a sequence  $k^N(u^N(\hat{t}_i))$  such that  $(k^N(u^N(\hat{t}_i)), \tilde{y}_i) \in T_{KY|U}^N(\epsilon_5)$ . If there exists a sequence, put out the corresponding index  $\hat{s}_i \oplus k_i^*$ . Otherwise, declare a decoding error. Finally, since receiver 1 knows  $k_i^*$  ( $k_i^* = g_f(\tilde{y}_{i-1})$ ), put out the corresponding  $\hat{s}_i$  from  $\hat{s}_i \oplus k_i^*$ .

2) *Proof of Achievability:* Note that the encoding and decoding scheme for Theorem 7 is exactly the same as that in [14], except that the transmitted message for receiver 1 is  $s \oplus k^*$ . Since receiver 1 knows  $k^*$ , the decoding scheme for Theorem 7 is in fact the same as that in [14]. Hence, we omit the proof of  $P_{e1} \leq \epsilon$  and  $P_{e2} \leq \epsilon$  here.

It remains to show that  $\lim_{N \rightarrow \infty} \Delta \geq R_e$ , see the following.

Since the confidential message  $S$  is encrypted by  $S \oplus K^*$ , the equivocation about  $S$  is equivalent to the equivocation about  $K^*$ . There are two ways for receiver 2 to obtain the secret key  $k^*$ . One way is that he tries to guess the  $k^*$  from its alphabet  $\mathcal{S}$ . The other way is that he tries to guess the feedback  $y^N$  ( $y^N$  is the output of channel 1 for the previous block, and  $k^* = g_f(y^N)$ ) from the conditional typical set  $T_{[Y|Z]}^N(\delta)$ , and this is because for a given  $z^N$  and sufficiently large  $N$ ,  $\Pr\{(y^N \notin T_{[Y|Z]}^N(\delta))\} \rightarrow 0$ . Note that there are  $2^{NH(Y|Z)}$  sequences  $y^N \in T_{[Y|Z]}^N(\delta)$  when  $N \rightarrow \infty$  and  $\delta \rightarrow 0$ . Therefore, the equivocation about  $S$  is  $\min\{\frac{\log \|\mathcal{S}\|}{N} = R_1, H(Y|Z)\}$ , and note that  $R_1 \geq R_e$  and  $H(Y|Z) \geq R_e$ , then  $\lim_{N \rightarrow \infty} \Delta \geq R_e$  is obtained.

The detail about the proof of  $\lim_{N \rightarrow \infty} \Delta \geq R_e$  is exactly the same as (82), and it is omitted here.

Thus,  $\lim_{N \rightarrow \infty} \Delta \geq R_e$  is proved, and the direct part of Theorem 7 is completed.

The proof of Theorem 7 is completed.

## REFERENCES

- [1] R. Ahlswede and N. Cai, "Transmission, Identification and Common Randomness Capacities for Wire-Tap Channels with Secure Feedback from the Decoder," book chapter in *General Theory of Information Transfer and Combinatorics*, LNCS 4123, pp. 258-275, Berlin: Springer-Verlag, 2006.
- [2] E. Ardestanizadeh, M. Franceschetti, T.Javidi and Y.Kim, "Wiretap channel with secure rate-limited feedback," *IEEE Trans Inf Theory*, vol. IT-55, no. 12, pp. 5353-5361, December 2009.
- [3] Y. Chen, A. J. Han Vinck, "Wiretap channel with side information," *IEEE Trans Inf Theory*, vol. IT-54, no. 1, pp. 395-402, January 2008.
- [4] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans Inf Theory*, vol. IT-24, no. 3, pp. 339-348, May 1978.
- [5] I. Csiszár and J. Körner, *Information Theory. Coding Theorems for Discrete Memoryless Systems*. London, U.K.: Academic, 1981.
- [6] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans Inf Theory*, vol. IT-29, no. 3, pp. 439-441, 1983.
- [7] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Control and Information Theory*, vol. 9, no. 1, pp. 19-31, 1980.
- [8] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans Inf Theory*, vol. IT-23, no. 1, pp. 60-64, January 1977.
- [9] N. V. Kuznetsov and B. S. Tsybakov, "Coding in memories with defective cells," *Problemy peredachi informatsii*, vol. 10, no. 2, pp. 52-60, 1974.
- [10] L. Lai, H. El Gamal and V. Poor, "The wiretap channel with feedback: encryption over the channel," *IEEE Trans Inf Theory*, vol. IT-54, pp. 5059-5067, 2008.
- [11] N. Merhav, "Shannon's secrecy system with informed receivers and its application to systematic coding for wiretapped channels," *IEEE Trans Inf Theory*, special issue on Information-Theoretic Security, vol. IT-54, no. 6, pp. 2723-2734, June 2008.
- [12] C. Mitrpant, A. J. Han Vinck and Y. Luo, "An Achievable Region for the Gaussian Wiretap Channel with Side Information," *IEEE Trans Inf Theory*, vol. IT-52, no. 5, pp. 2181-2190, 2006.
- [13] C. E. Shannon, "Channels with side information at the transmitter," *IBM Journal Research and Development*, vol. 2, pp. 289-293, October 1958.
- [14] Y. Steinberg, "Coding for the degraded broadcast channel with random parameters, with causal and noncausal side information," *IEEE Trans Inf Theory*, vol. IT-51, no. 8, pp. 2867-2877, August 2005.
- [15] A. D. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355-1387, 1975.